

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM nucleic - nucleic search, using sw model

Run on: August 4, 2003, 21:01:54 ; Search time 7372.88 Seconds

(without alignments)  
11491.278 Million cell updates/sec

Title: US-09-936-271B-13\_COPY\_9500\_11570

Perfect score: 2071

Sequence: 1 aaacagaccacaaactctct.....cccaagaataactgagaag 2071

Scoring table:

IDENTITY\_NMC  
Gapop 10.0 , Gapept 1.0

Searched: 2888711 seqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenEmbl:*
1: gb_da:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_ph:*
24: em_pl:*
25: em_ro:*
26: em_sts:*
27: em_un:*
28: em_vi:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_jny:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pla:*
35: em_htg_rod:*
36: em_htg_man:*
37: em_htg_vrl:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2071	100.0	11570	9 AF135028	AF135028 Homo sapi
2	2071	100.0	107487	9 AC011483	AC011483 Homo sapi
3	2071	100.0	230000	9 AF243527	AF243527 Homo sapi
4	1990	96.1	217346	2 AC027602	AC027602 Homo sapi
5	1686.8	81.4	200792	2 AC130782	AC130782 Pan trogl
6	482	23.3	586	6 AX331407	AX331407 Sequence
7	482	23.3	586	6 AX331824	AX331824 Sequence
8	482	23.3	586	11 G41906	G41906 SHG-56840
9	482	23.3	735	6 AX429955	AX429955 Sequence
10	482	23.3	738	6 BD139877	BD139877 Compounds
11	482	23.3	1370	9 AY279380	AY279380 Homo sapi
12	482	23.3	1381	6 BD107879	BD107879 36 human
13	482	23.3	1387	9 AF168768	AF168768 Homo sapi
14	482	23.3	1438	9 AY279381	AY279381 Homo sapi
15	482	23.3	1439	6 BD107880	BD107880 36 human
16	482	23.3	1499	6 BD005362	BD005362 Protease.
17	482	23.3	1516	6 BD107865	BD107865 36 human
18	482	23.3	1536	6 BC008036	BC008036 Homo sapi
19	482	23.3	1570	6 AR252580	AR252580 Sequence
20	482	23.3	1570	6 AX080829	AX080829 Sequence
21	482	23.3	1570	6 AX403421	AX403421 Sequence
22	482	23.3	1570	6 AX464322	AX464322 Sequence
23	476	23.0	1476	6 AR078184	AR078184 Sequence
24	476	23.0	1476	6 AR137506	AR137506 Sequence
25	476	23.0	1476	6 AR242358	AR242358 Sequence
26	476	23.0	1476	6 BD082659	BD082659 Keratinoc
27	460	22.2	1504	6 AR263823	AR263823 Sequence
28	449	21.7	673	6 AX067353	AX067353 Sequence
29	270.4	13.1	57729	9 AL449284	AL449284 Human DNA
30	264.2	12.8	42107	9 AC004409	AC004409 Homo sapi
31	261.2	12.6	129048	9 AC020931	AC020931 Homo sapi
32	260	12.6	120997	9 AC013449	AC013449 Homo sapi
33	255.8	12.4	23615	2 AP000933	AP000933 Homo sapi
34	252.2	12.2	118085	9 AL136115	AL136115 Human DNA
35	250.8	12.1	172945	9 AC007220	AC007220 Homo sapi
36	249.8	12.1	160915	2 AP001084	AP001084 Homo sapi
37	249.8	12.1	165420	9 AC091588	AC091588 Homo sapi
38	249.8	12.1	171978	2 AC009669	AC009669 Homo sapi
39	249.8	12.1	204504	2 AC022487	AC022487 Homo sapi
40	249.8	12.0	208310	9 AC091043	AC091043 Homo sapi
41	249	12.0	126205	2 AP002344	AP002344 Homo sapi
42	247.4	11.9	190982	2 AC025285	AC025285 Homo sapi
43	246.8	11.9	160858	9 AC019216	AC019216 Homo sapi
44	246	11.9	152659	9 AL591503	AL591503 Human DNA
45	243.2	11.7	40395	9 AC099491	AC099491 Homo sapi

## ALIGNMENTS

RESULT 1	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
AF135028	AF135028	Homo sapiens kallikrein-like protein 2 KLK-L2 gene, complete cds.	AF135028	AF135028.1	GI:4589282	Homo sapiens (human)	Homo sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	Yousef, G.M. and Diamandis, E.P.	The new kallikrein-like gene, KLK-L2. Molecular characterization, mapping, tissue expression, and hormonal regulation



D	b		11000	TCCATCATATGGGCAATCTCTBGGGTCTCTCATGTCTCCTTTGCCCCAATTG6CCACATCTT	11058
OY			1561	GCCCTCTCTCATGGCCCCCTTTCTCTCTCTCGAGGGATGTCCTGGGGGCCCTGTGTCTTGCA	1620
D	b		11060	GCCTCTCTCATGGCCCCCTTTCTCTCTCTCGAGGGATGTCCTGGGGGCCCTGTGTCTTGCA	11119
OY			1621	ATGGCTCTCCCTGCAGAGGACTGTGTCTCTGGGAGATTACCCTTGTGCCGGCCACAGAC	1680
D	b		11120	ATGGCTCTCCCTGCAGAGGACTGTGTCTCTGGGAGATTACCCTTGTGCCGGCCACAGAC	11179
OY			1681	CGGGTGTCTACAGAACTCTGAAGTCCAAAGTGATCCAGGAACCATCCAGGCCA	1740
D	b		11180	CGGGTGTCTACAGAACTCTGAAGTCCAAAGTGATCCAGGAACCATCCAGGCCA	11239
OY			1741	ACTCTGAGTCATCCAGAGACTCAGACACCGGATCCACCCTGCTGCAGGAGCACGCC	1800
D	b		11240	ACTCTGAGTCATCCAGAGACTCAGACACCGGATCCACCCTGCTGCAGGAGCACGCC	11239
OY			1801	TGACATCTCTTTAGACACCTCATCTCTCTCCAGAGATGTTGAGATGTTATCTCTCAG	1860
D	b		11300	TGACATCTCTTTAGACACCTCATCTCTCTCCAGAGATGTTGAGATGTTATCTCTCAG	11359
OY			1861	CCCCGTGACCCCATGTCTCTGAGCTCAGGGTGTCTCTCCCATCTGGGGCTGACCGTCT	1920
D	b		11360	CCCCGTGACCCCATGTCTCTGAGCTCAGGGTGTCTCTCCCATCTGGGGCTGACCGTCT	11419
OY			1921	CTCTTAGTTGAACCCCTGGGACAATTTCCAAAACGTCTCAGGGCCGGGGTTGCGTCTCA	1980
D	b		11420	CTCTTAGTTGAACCCCTGGGACAATTTCCAAAACGTCTCAGGGCCGGGGTTGCGTCTCA	11479
OY			1981	ATCTCCCTGGGGCACTTTCATCTCTCAAAGCTCAGGGCCCATCTCTCTGACGCTCTGAC	2040
D	b		11480	ATCTCCCTGGGGCACTTTCATCTCTCAAAGCTCAGGGCCCATCTCTCTGACGCTCTGAC	11539
OY			2041	CCAAATTAGTCCAGAAATAAAGTACAGAG	2071
D	b		11540	CCAAATTAGTCCAGAAATAAAGTACAGAG	11570
RESULT 2					
LOCUS			AC011483/c		
DEFINITION			Homo sapiens chromosome 19 clone CTR-147c22, complete sequence.		
ACCESSION			AC011483		
VERSION			AC011483.7		
KEYWORDS			HTG.		
SOURCE			Homo sapiens (human)		
ORGANISM			Homo sapiens		
REFERENCE			Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS			Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
TITLE			1 (bases 1 to 107487)		
JOURNAL			DOE Joint Genome Institute and Stanford Human Genome Center.		
REFERENCE			Direct Submission		
AUTHORS			Unpublished		
TITLE			2 (bases 1 to 107487)		
JOURNAL			DOE Joint Genome Institute.		
REFERENCE			Direct Submission		
AUTHORS			Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint		
TITLE			Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA		
JOURNAL			3 (bases 1 to 107487)		
REFERENCE			DOE Joint Genome Institute and Stanford Human Genome Center.		
AUTHORS			Direct Submission		
TITLE			Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell		
JOURNAL			Drive, Walnut Creek, CA 94598, USA		
COMMENT			On Jul 1, 2002 this sequence version replaced gi:14971176.		
			Draft Sequence Produced by DOE Joint Genome Institute		
			www.jgi.doe.gov		
			Finishing Completed at Stanford Human Genome Center		
			www.sngc.stanford.edu		
			Quality: Phrap Quality >=40 99.6% of Sequence;		
			Estimated Total Number of Errors Is 0.6.		
			NOTE: Shatter Libraries failed to resolve dinucleotide repeat.		

FEATURES	Unsure number of repeat copies	64998-65494	Forced join	65015
source	Location/Qualifiers			
	1. .107487			
	/organism="Homo sapiens"			
	/mol_type="genomic DNA"			
	/db_xref="taxon:9606"			
	/chromosome="19"			
	/clone="CtB-147C2"			
	/note="NOTE: Shatter libraries failed to resolve dinucleotide repeat. Unsure number of repeat copies 64998-65494. Forced join 65015."			
BASE COUNT	26044 a 27304 c 25728 g 28411 t			
ORIGIN				
Query Match	100.0%; Score 2071; DB 9; Length 107487;			
Best Local Similarity	100.0%; Pred. No. 0;			
Matches 2071; Conservative	0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 AAACAGCCAAAACTTCCTGCGCGGTGACCTCATGTTCCCAAGTGAAGACAGCAA	60		
Db	50720 AAACAGACCCAAAACTTCCTGCGCGGTGACCTCATGTTCCCAAGTGAAGACAGCAA	50661		
QY	61 TAAAGACATGATTAATATGTAGTAAATTTAAAAAATTTATGCCGGGTGTGTGG	120		
Db	50660 TAAAGACATGATTAATATGTAGTAAATTTAAAAAATTTATGCCGGGTGTGTGG	50601		
QY	121 CTTTCACCTGTAGTTCACGCTACTTGGGAGGCTGAGGTGGAGAAATTCCTTGAGCCCAA	180		
Db	50600 CTTTCACCTGTAGTTCACGCTACTTGGGAGGCTGAGGTGGAGAAATTCCTTGAGCCCAA	50541		
QY	181 CGTTTGAGGCTGCGGTAAAGCCATGACTGACCTGCTCCTCCAGACGAGCCTGGGTGA	240		
Db	50540 CGTTTGAGGCTGCGGTAAAGCCATGACTGACCTGCTCCTCCAGACGAGCCTGGGTGA	50481		
QY	241 CAAAGCAAGACGTTTTGTGTCAGAAAGAAAAAAGACACAGGAGGAGAGAGAG	300		
Db	50480 CAAAGCAAGACGTTTTGTGTCAGAAAGAAAAAAGACACAGGAGGAGAGAGAG	50421		
QY	301 AAAGGAAGGAAGGAGAGAAAGAAAGGAAGGAAGGAAGAAAGGAAGGAAGGA	360		
Db	50420 AAAGGAAGGAAGGAGAGAAAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGA	50361		
QY	361 GAAAGAAAGGAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA	420		
Db	50360 GAAAGAAAGGAGAAAGAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA	50301		
QY	421 GAGGAGGAAAGAAAGAGAAAGAAAGGAAAAAATGACTGTTGAAGAGCAGTGAT	480		
Db	50300 GAGGAGGAAAGAAAGAGAAAGAAAGGAAAAAATGACTGTTGAAGAGCAGTGAT	50241		
QY	481 TATATATGAGGAGGTAATATTAGGAGAGGTATGGGGAATTTGAAGACAGAAACCAATTTG	540		
Db	50240 TATATATGAGGAGGTAATATTAGGAGAGGTATGGGGAATTTGAAGACAGAAACCAATTTG	50181		
QY	541 TCCAAAGCAATGATTTCTATTGGAGGTGAATTTGCCCTAGAAAGACACTGGCAATACCA	600		
Db	50180 TCCAAAGCAATGATTTCTATTGGAGGTGAATTTGCCCTAGAAAGACACTGGCAATACCA	50121		
QY	601 GGAGACATTTTGGTTGTCAACACTATATGAGGGGCAATTTACTGCCAATATGATAGA	660		
Db	50120 GGAGACATTTTGGTTGTCAACACTATATGAGGGGCAATTTACTGCCAATATGATAGA	50061		
QY	661 TGGCAACTGTGCTGTTCAACATGCTATGATGACACAGGGAGGCTCCACAACAACCAAT	720		
Db	50060 TGGCAACTGTGCTGTTCAACATGCTATGATGACACAGGGAGGCTCCACAACAACCAAT	50001		
QY	721 ATCCAGCTTATGATGCCACAGTGCACAGTTCAGAGGAACCTTATCCAGGGGCTGAGAAC	780		
Db	50000 ATCCAGCTTATGATGCCACAGTGCACAGTTCAGAGGAACCTTATCCAGGGGCTGAGAAC	49941		
QY	781 CGTATTTTTCAGAGGAGGATTAAGATGGTTGTGGAGAAATGGGAGAGGAAGTGT	840		

Db 49940 CGTATTTTTCGAGAGGAGGTATTAAGATGGGTTGTGTGAGAAATGGGGAAGAGTGT 49881  
 Oy 841 GTGTCCAGTAAAGAAATAGCCCTGCACAGCTTGAGGGGAGATGAGAGAAAGGA 900  
 Db 49880 GTGTCCAGTAAAGAAATAGCCCTGCACAGCTTGAGGGGAGATGAGAGAAAGGA 49821  
 Oy 901 GGGGAGAGATACAGATGAGGGAGACAGGCTGGAACAGAAAGTAGACGAAAGATTGGA 960  
 Db 49820 GGGGAGAGATACAGATGAGGGAGACAGGCTGGAACAGAAAGTAGAGCAAGATTGGA 49761  
 Oy 961 GATGTGAGAGAGAGGCTCACAGACCCCGGAAATGATGTGTGACAAACAGATCTGG 1020  
 Db 49760 GATGTGAGAGAGAGGCTCACAGACCCCGGAAATGATGTGTGACAAACAGATCTGG 49701  
 Oy 1021 AAGAGAGATGAGAGTGAAGTGAACAATGGGCTTAAGGTTGAATCTTGAGGCCAG 1080  
 Db 49700 AAGAGAGATGAGAGTGAAGTGAACAATGGGCTTAAGGTTGAATCTTGAGGCCAG 49641  
 Oy 1081 CATGTGAGCTACAGGCTGTAAATCCCAACACTTTGGAGGCTGAGTGGCGAATCACTGA 1140  
 Db 49640 CATGTGAGCTACAGGCTGTAAATCCCAACACTTTGGAGGCTGAGTGGCGAATCACTGA 49581  
 Oy 1141 GGGCAGAGTTCGAGACCAAGCTGGCCAAACATGCTGAACCCCTCTCTACAAAAAAT 1200  
 Db 49580 GGGCAGAGTTCGAGACCAAGCTGGCCAAACATGCTGAACCCCTCTCTACAAAAAAT 49521  
 Oy 1201 ACAAATAATAGCCGGGTGTGTGTGAGACACTGTAGTCAACAGTACTTGAGAGCTGA 1260  
 Db 49520 ACAAATAATAGCCGGGTGTGTGTGAGACACTGTAGTCAACAGTACTTGAGAGCTGA 49461  
 Oy 1261 GGCAGAGAAATGCTTGAACCCGGGAGATGAGAGCTGCAAGTGAAGTCAAGCTCACT 1320  
 Db 49460 GGCAGAGAAATGCTTGAACCCGGGAGATGAGAGCTGCAAGTGAAGTCAAGCTCACT 49401  
 Oy 1321 GGGCTCCAACTGGGCAACAGAGTAAAGTCCATCTCAAAAAAAGCTGATTTGG 1380  
 Db 49400 GGGCTCCAACTGGGCAACAGAGTAAAGTCCATCTCAAAAAAAGCTGATTTGG 49341  
 Oy 1381 AGTGAATATTAATTAACATCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440  
 Db 49340 AGTGAATATTAATTAACATCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 49281  
 Oy 1441 TTTTCTGATTTCTTCACTCTGTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500  
 Db 49280 TTTTCTGATTTCTTCACTCTGTAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 49221  
 Oy 1501 TCCATCTAAGGCACTCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560  
 Db 49220 TCCATCTAAGGCACTCTGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 49161  
 Oy 1561 GCGTCTCATGCCCCCTTCTCTCTCTGACAGGATGATCTGGGGGCTGTGTCTGCA 1620  
 Db 49160 GCGTCTCATGCCCCCTTCTCTCTCTGACAGGATGATCTGGGGGCTGTGTCTGCA 49101  
 Oy 1621 ATGCTCTCCCTGACAGGACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680  
 Db 49100 ATGCTCTCCCTGACAGGACTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 49041  
 Oy 1681 GGGTGTCTAGCAACCTCTGCAAGTTCACCAAGTGTCCAGGAACATCCAGGGCA 1740  
 Db 49040 GGGTGTCTAGCAACCTCTGCAAGTTCACCAAGTGTCCAGGAACATCCAGGGCA 48981  
 Oy 1741 ACTCTGAGTATCCAGGACTCAGACACCGGCACTCCCACTCTGAGGAGACGCC 1800  
 Db 48980 ACTCTGAGTATCCAGGACTCAGACACCGGCACTCCCACTCTGAGGAGACGCC 48921  
 Oy 1801 TGACACTCTCTCAGACCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860  
 Db 48920 TGACACTCTCTCAGACCTCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 48861  
 Oy 1861 CCCGAGACCATGCT 1920  
 Db 48860 CCCGAGACCATGCT 48801

Oy 1921 CTCTCTAGTTGACCCCTGGGAACAATTTCCAAAACCTGTCAGGGCGGGGTTGGCTCTCA 1980  
 Db 48800 CTCTCTAGTTGACCCCTGGGAACAATTTCCAAAACCTGTCAGGGCGGGGTTGGCTCTCA 48741  
 Oy 1981 ATCTCCCTGGGGCACTTTCATCTCTCAAGCTGAGGGCCCATCTCTCTCTCTCTCTGAC 2040  
 Db 48740 ATCTCCCTGGGGCACTTTCATCTCTCAAGCTGAGGGCCCATCTCTCTCTCTCTGAC 48681  
 Oy 2041 CCAATTTAGTCCGAAATTAACAGAGAG 2071  
 Db 48680 CCAATTTAGTCCGAAATTAACAGAGAG 48650

RESULT 3  
 AF243527/c  
 LOCUS Homo sapiens serine protease gene cluster, complete sequence.  
 DEFINITION  
 ACCESSION AF243527  
 VERSION AF243527.1 GI:11244757  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE  
 Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,  
 Moss, P., Paepert, B. and Wang, K.  
 Sequencing and expression analysis of the serine protease gene  
 cluster located in chromosome 19q13 region  
 JOURNAL Gene 257 (1), 119-130 (2000)  
 MEDLINE 20510030  
 PUBMED 11054574  
 REFERENCE  
 2 (bases 1 to 230000)  
 Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuaig, J.,  
 Moss, P., Paepert, B. and Wang, K.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,  
 Bothell, WA 98021, USA  
 FEATURES  
 source  
 1..230000  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /map="19q13"  
 /gene="KLK1"  
 complement(1..6963..>11517)  
 /gene="KLK1"  
 complement(1..7118..7668..7804..7923..8212,  
 9482..9641..11472..>11517)  
 /gene="KLK1"  
 /product="renal kallikrein"  
 complement(1..6963..7118..7668..7804..7923..8212,  
 9482..9641..11472..11517)  
 /gene="KLK1"  
 /note="serine protease"  
 /codon\_start=1  
 /product="renal kallikrein"  
 /protein\_id="AAG33353.1"  
 /db\_xref="GI:11244758"  
 /translation="MNPFLVCLASLSTGTAAPPIQSIYGVHCEQHSOPWQALAH  
 FSTIQGGILVHKRWLVLAHCISDNTOLMIGRNLTPDDENTAQFVHSEFPHGN  
 MSLENNHROADEVYSHDLMLRLTEPADTITDAVKVLEPTEPEVSGSTLASGWS  
 IEPENFSPDLOGVDLKILPNDECKKAHVQKVDPMVLELGGKDKTCVDSGGPL  
 MCDVLOGVTSNGVPGCTPNKPSVAIVLVSVKWIETLAEINS"  
 complement(1..13552..13704..14377..14530..14678..14917,  
 15416..15569..19204..>19246)  
 /product="ACO protease"  
 complement(1..13552..13704..14377..14530..14678..14917,  
 15416..15569..19204..19246)  
 /note="serine protease"  
 /codon\_start=1  
 /product="ACO protease"

Query Match	100.0%; Score 2071; DB 9; Length 230000;
Best Local Similarity	100.0%; Pred. No. 0;
mRNA	<p>/translation="MWVLCALITALLIGVTDARSDDSSSRITNGSDCDMTHTPMAQAL  ILRPMDQCGAVLTHPPDWMLITAAHCRRKRVFVRLGHSLTSVPESSGGOMQGVKSJPH  PYSRPHGSHNDMLTIKLNRRIRPPIKDVPIVSSGHSGATKCLSVSGMGTSPQVHF  PVYLQCLNIVSLSCQKCEDAVPROIDMTMPCGAKGACGSDSQGDSGGPVCNYSLOGL  VSMGVPICARPRPVCVYNTICKFTFKWIDTIOANS"  complement(join(&lt;146834 . 146966,149628 . 149764,  151186 . 151433,155052 . 155208,155948 . &gt;155987))  /product="protease M"  complement(join(146834 . 146986,149628 . 149764,  151186 . 151433,155052 . 155208,155948 . 155987))  /notes="serine protease; also called neurosin or zyme"  /codon_start=1  /product="protease M"  /protein_id="AAG33359.1"  /db_xref="GI:11244764"  /translation="MKRLMVLSLITAAAEKQKIVHGPGCDKTSHPQALTSCH  LLCGVLIHPMLVLTAAHCRRKPNQVFLGRKLNLRQRESQDSSVYAAVTHPYTDAS  HPODITMLRLARPAKLSLEIOPLEPERSANNTSCHILGWKTADDPDTIOCATI  HLVSEDEHAYPGQITQNMCLACGDEKXGKSDCGSDGSLVCGDHLRGLVSGMNIPLC  GSEKRPGYVATCRFTYNTIKTIOAK"  complement(&lt;165420 . &gt;170283)  /genes="KIK7"  /notes="synonym: SCCP"  complement(join(&lt;165420 . 165575,167672 . 167808,  168124 . 168371,169651 . 169798,170211 . &gt;170283))  /genes="KIK7"  /product="stratum corneum chymolytic enzyme"  complement(join(165420 . 165575,167672 . 167808,  168124 . 168371,169651 . 169798,170211 . 170283))  /genes="KIK7"  /notes="serine protease"  /codon_start=1  /product="stratum corneum chymolytic enzyme"  /protein_id="AAG33360.1"  /db_xref="GI:11244765"  /translation="MARSLPLPQILLSTALETAEGEAQGDKIIDGAPCARSGHPMQ  VALISGNDLHCGGLVNERWVLTAAHCMMBYTVHLSDTLGRRAQRIKASGFRHP  GISTQTHVDMLKVLINSQARLSVYKVRRLPSRCEPFTGTVSGMGTTSPTVTP  SOLMGVLDLSPDCTKVYKDLLENSMLCAGIPDSKNAACNGDSGGPLVCRGTLQGL  VSMGFPCGPPDPVYVYQVCKFTKIMDPKKHR"  complement(join(&lt;183943 . 184098,185635 . 185768,  187865 . 188127,188293 . 188452,188967 . &gt;189036))  /product="neutropsin"  complement(join(183943 . 184098,185635 . 185768,  187865 . 188127,188293 . 188452,188967 . 189036))  /notes="serine protease; also called ovasin"  /codon_start=1  /product="neutropsin"  /protein_id="AAG33361.1"  /db_xref="GI:11244766"  /translation="MGRPRPAAKTWMLTLLGGAMGHSRAQEDKVLGHEGCPHSO  PMAQLFQGOGLCGGVYGVGNVLTAAHCCKRPYTVRLGHSLOKNDGDEPELTPVQ  SLPHECVSDVEDHNDMLDLQLEKQDSLSGSKRPLSLAHCTQPOQKCTVSQWGV  TSPENFPTDLNCAEVKIFPQKCEDAVPGQITGMVCAQSGSKADDPQDSSGPLVC  DEALOGITSMGSDPCGSRDGPVYNTICRIYDWKILIGSG"  complement(join(&lt;190980 . 191129,191573 . 191709,  194334 . 194589,197048 . 197204,197370 . &gt;197412))  /product="kallikrein-like 3"  complement(join(190980 . 191129,191573 . 191709,  194334 . 194589,197048 . 197204,197370 . 197412))  /notes="serine protease"  /codon_start=1  /product="kallikrein-like 3"  /protein_id="AAG33362.1"  /db_xref="GI:11244767"  /translation="MKILGICALSLSLAGHGMADTRAIGAECHRPNSQPMQAGLFIHLT  RLFCAATLISRMILTAAHCRRPYLWVLSGHLIMKMGEGEQLRVYDFPFGFNKD  ISANNDHIDMLRLPRLARSLPAVQPIVNTQVCSGMOLISGCAVSSPALPVP  TLOCAITSLIENKICMAVPGHISDSMLCALMEGSGSCQGDGSLVNGTLGACV  SGGAFCPSRRPRAVYTSVCHLMDIDETMEN"</p>
CDS	
mRNA	
CDS	
mRNA	
CDS	
gene	
mRNA	
CDS	

Matches 2071: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 AAACAGACCAAACTTCTCCGCGTGACCTCATGTGTTCCCAAGTGAAGACAGCA 60  
133048 AAACAGACCAAACTTCTCCGCGTGACCTCATGTGTTCCCAAGTGAAGACAGCA 132989  
OY 61 TAAAGATAGATATAATATGTATATTAATAAAAAAAAAAAATTTAGCCGGGTGTGTG 120  
132988 TAAAGATAGATATAATATGTATATTAATAAAAAAAAAAAATTTAGCCGGGTGTGTG 132929  
OY 121 CTGTGACCTGATGTCCAGCTACTTGGGAGGCTGAGGTGGAGAAATTCCTTGAGCCCA 180  
132928 CTGTGACCTGATGTCCAGCTACTTGGGAGGCTGAGGTGGAGAAATTCCTTGAGCCCA 132869  
OY 181 CGTTGAGGCTGCGGTAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
132868 CGTTGAGGCTGCGGTAGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 132809  
OY 241 CAAACGAAGCGTTTGTCTAGAAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300  
132808 CAAACGAAGCGTTTGTCTAGAAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 132749  
OY 301 AAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 360  
132748 AAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 132689  
OY 361 GAAAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 420  
132688 GAAAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 132629  
OY 421 GAGGAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 480  
132628 GAGGAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 132569  
OY 481 TATTTAGAGGCTAATATAGGAGGTATGGGAAATGAAACAGAAACCAATTTAG 540  
132568 TATTTAGAGGCTAATATAGGAGGTATGGGAAATGAAACAGAAACCAATTTAG 132509  
OY 541 TCCAGCGAATGATTTCTATTTGGAGATGATTTCTGCCCCCTAGAGACACTGGCAATACCA 600  
132508 TCCAGCGAATGATTTCTATTTGGAGATGATTTCTGCCCCCTAGAGACACTGGCAATACCA 132449  
OY 601 GGAGACATTTTGTGTGTCACACTATATGAGGGGCAATTTACTGGCACTAATGATAGA 660  
132448 GGAGACATTTTGTGTGTCACACTATATGAGGGGCAATTTACTGGCACTAATGATAGA 132389  
OY 661 TGGCAAGTGTGCTGTCACACTATGATGATGCAACGGGCGCTCCCAACAAACCAAT 720  
132388 TGGCAAGTGTGCTGTCACACTATGATGATGCAACGGGCGCTCCCAACAAACCAAT 132329  
OY 721 ATCCAGCTTCAGATGCCACAGTGGCCAGATCGAGGAACCTTCATCCAGGGGCTGAGAAC 780  
132328 ATCCAGCTTCAGATGCCACAGTGGCCAGATCGAGGAACCTTCATCCAGGGGCTGAGAAC 132269  
OY 781 CGTATTTTTCAGAAAGGAGATATAAGATGGGTGTGTGAGAAATGGGAGAGAGAGTGT 840  
132268 CGTATTTTTCAGAAAGGAGATATAAGATGGGTGTGTGAGAAATGGGAGAGAGTGT 132209  
OY 841 GGTGTCAGTAAGAAATTAAGGCTGCAACAGCTGGAGGGGAGAGAGAGAGAGAGAGAGAG 900  
132208 GGTGTCAGTAAGAAATTAAGGCTGCAACAGCTGGAGGGGAGAGAGAGAGAGAGAGAGAG 132149  
OY 901 GGCGGAGATACACAGTAGGAG 960  
132148 GGCGGAGATACACAGTAGGAG 132089  
OY 961 GATGTGAG 1020  
132088 GATGTGAG 132029  
OY 1021 AAGAGAGAGATGAGTGAAG 1080  
132028 AAGAGAGAGATGAGTGAAG 131969

OY 1081 CATGTGGCTCAGCGCTGTAAATCCCAACACTTTGGAGGCTGAGGTGGGCAATCATTTGA 1140  
131968 CATGTGGCTCAGCGCTGTAAATCCCAACACTTTGGAGGCTGAGGTGGGCAATCATTTGA 131909  
OY 1141 GGCCAGAGATTCGAGACAGCAGCTGCGCAACATGTTGAACCCGCTCTCAAAAAAAAT 1200  
131908 GGCCAGAGATTCGAGACAGCAGCTGCGCAACATGTTGAACCCGCTCTCTCAAAAAAAAT 131849  
OY 1201 ACAAATAATTAAGCGGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260  
131848 ACAAATAATTAAGCGGCTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 131789  
OY 1261 GGCAAGAGATTCGTTGAACCCGAGATGAGAGCTGCACTGAGCTGAGCTGAGCTGAGCTGAG 1320  
131788 GGCAAGAGATTCGTTGAACCCGAGATGAGAGCTGCACTGAGCTGAGCTGAGCTGAGCTGAG 131729  
OY 1321 GCGCTCCAGCTGGGCAACAGATGAAGCTCATCTCAAAAAAAGCTGATTTGG 1380  
131728 GCGCTCCAGCTGGGCAACAGATGAAGCTCATCTCAAAAAAAGCTGATTTGG 131669  
OY 1381 AGTGAATATTAATTAACATTCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440  
131668 AGTGAATATTAATTAACATTCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 131609  
OY 1441 TTTTCTGATTTCTTCACT 1500  
131608 TTTTCTGATTTCTTCACT 131549  
OY 1501 TCCATCTATGGGCACTCTGCGGCT 1560  
131548 TCCATCTATGGGCACTCTGCGGCT 131489  
OY 1561 GCGCTCTCATGCCCCCTTCT 1620  
131488 GCGCTCTCATGCCCCCTTCT 131429  
OY 1621 ATGCTCTCTCTGAGGACTCTGCT 1680  
131428 ATGCTCTCTCTGAGGACTCTGCT 131369  
OY 1681 GGGGTGTCTACAGCAACCTCTGCAAGTTCACCAAGTGTGATGATGATGATGATGATGATGATG 1740  
131368 GGGGTGTCTACAGCAACCTCTGCAAGTTCACCAAGTGTGATGATGATGATGATGATGATGATG 131309  
OY 1741 ACTCTGAGTCAATCCAGAGACTCAGACACCGGCAATCCCACTGCTCAGAGGACAGCC 1800  
131308 ACTCTGAGTCAATCCAGAGACTCAGACACCGGCAATCCCACTGCTCAGAGGACAGCC 131249  
OY 1801 TGACACTCTCTTCAACACCTCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1860  
131248 TGACACTCTCTTCAACACCTCATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 131189  
OY 1861 CCGCTGACCCCATGCT 1920  
131188 CCGCTGACCCCATGCT 131129  
OY 1921 CTCTCTAGTTGAACCTCTGGAACAATTTCCAAACTGTCCAGAGGCGGGGCTGCTCTCA 1980  
131128 CTCTCTAGTTGAACCTCTGGAACAATTTCCAAACTGTCCAGAGGCGGGGCTGCTCTCA 131069  
OY 1981 ATCTCCCTGGGCACTTTTCT 2040  
131068 ATCTCCCTGGGCACTTTTCT 131009  
OY 2041 CCAAAATTTAGTCCCAAGAAATTAAGTGAAG 2071  
131008 CCAAAATTTAGTCCCAAGAAATTAAGTGAAG 130978

RESULT 4  
AC027602/c AC027602 217346 bp DNA linear HTG 24-AUG-2002  
LOCUS

```

DEFINITION Homo sapiens chromosome 19 clone.RP11-795B6 map 19, WORKING DRAFT
SEQUENCE 12 unordered pieces.
ACCESSION AC027602.4 GI:11178143
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 19, clone RP11-795B6
JOURNAL Unpublished
REFERENCE
AUTHORS 2 (bases 1 to 217346)
1 (bases 1 to 217346)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 19, clone RP11-795B6
Unpublished
2 (bases 1 to 217346)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouckhalter,B., Brown,A., Burkett,G.,
Campoliano,A., Castle,A., Choehel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
McDermid,J., Menes,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tessier,S., Theodore,J., Tirrell,A., Travers,M., Triggilio,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

JOURNAL Direct Submission
TITLE Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS 3 (bases 1 to 217346)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,
Bouckhalter,B., Brown,A., Burkett,G., Campoliano,A., Castle,A.,
Choehel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,
Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., Laroque,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,
Macdonald,P., Margulis,N., McCarthy,M., McEwan,P., McKernan,K.,
McPheeters,R., Meldrum,J., Menes,L., Mihova,T., Mlenga,V.,
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,
O'Donnell,P., O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tessier,S., Theodore,J.,
Tirrell,A., Travers,M., Triggilio,J., Vassiliev,H., Viel,R., Vo,A.,
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
Zimmer,A. and Zody,M.

JOURNAL Direct Submission
TITLE Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT On Nov 15, 2000 this sequence version replaced gi:11136831.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L9166

```

```

Center Clone name: 795.B.6
----- Summary Statistics
Sequencing vector: M13, M7815; 31% of reads
Sequencing vector: Plasmid; n/a; 69% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 210748 bases at least Q40
Consensus quality: 213655 bases at least Q30
Consensus quality: 215058 bases at least Q20
Insert size: 194000; agarose-fp
Quality coverage: 11.9 in Q20 bases; agarose-fp
Quality coverage: 10.7 in Q20.
NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
8150 8149: contig of 8149 bp in length
8250 8249: gap of 100 bp
8250 9592: contig of 1343 bp in length
9593 9592: gap of 100 bp
9593 10733: contig of 1041 bp in length
10734 10833: gap of 100 bp
10834 10833: gap of 2686 bp in length
13520 13619: gap of 100 bp
13620 17510: contig of 3891 bp in length
17511 17610: gap of 100 bp
17611 24602: contig of 6992 bp in length
24603 24702: gap of 100 bp
24703 35434: contig of 10732 bp in length
35435 35535: gap of 100 bp
35535 124474: contig of 88940 bp in length
124475 124574: gap of 100 bp
124575 134664: contig of 10090 bp in length
134665 134764: gap of 100 bp
134765 162343: contig of 27579 bp in length
162344 208917: contig of 46474 bp in length
208918 209017: gap of 100 bp
209018 217346: contig of 8329 bp in length.
FEATURES
source
Location/Qualifiers
1. 217346
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19"
/clone="RP11-795B6"
/clone_1lb="RPC1-11 Human Male BAC"
1. 8149
/clone="assembly_fragment"
vector_end:Sp6
vector_side:left"
8250..9592
/clone="assembly_fragment"
9693..10733
/clone="assembly_fragment"
10834..13519
/clone="assembly_fragment"
13620..17510
/clone="assembly_fragment"
17611..24602
/clone="assembly_fragment"
24703..35434
/clone="assembly_fragment"
35535..124474
/clone="assembly_fragment"
124575..134664
/clone="assembly_fragment"

```

```
misc_feature 134765..162343
                /note="assembly-fragment"
misc_feature 162444..208917
                /note="assembly-fragment"
misc_feature 209018..217346
                /note="assembly-fragment"
                clone_end:77
                vector_side:right"
BASE COUNT 55109 a 54154 c 53923 g 53053 t 1107 others
ORIGIN
Query Match 96.1%; Score 1990; DB 2; Length 217346;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 2045; Conservative 0; Mismatches 5; Indels 22; Gaps 3;
QY 1 AAACAGACCAAAACCTCCGCGCGCTGAGACCTCATGTGCCCAAGTGGAGACAGCA 60
Db 194902 AAACAGACCAAAACCTCCGCGCGCTGAGACCTCATGTGCCCAAGTGGAGACAGCA 194843
QY 61 TAAAGATAGATTAATATGTAGTAATTAATAAAAAAAAAAAAAAAAAATTAGCCGGGTGTGTG 120
Db 194842 TAAAGATAGATTAATATGTAGTAATTT-TAAAAAATAATTTAGCCGGGTGTGTG 194785
QY 121 CTTGACCTGTAGTTCACACTACTTGGAGGCTGAGAGTGGAGAAATTGCTTGAGCCCAA 180
Db 194784 CTTGACCTGTAGTTCACACTACTTGGAGGCTGAGAGTGGAGAAATTGCTTGAGCCCAA 194725
QY 181 CTTTGAAGCTCGCGTAAACCATGACTGCACTGCTCAGTCCAGACAGCAGCTGGGTGA 240
Db 194724 CTTTGAAGCTCGCGTAAACCATGACTGCACTGCTCAGTCCAGACAGCAGCTGGGTGA 194665
QY 241 CAAGCAAGACGTTTTTGTGCAAGAAAAAAGAGACGAGAGGAGAGAGAGAG 300
Db 194664 CAAGCAAGACGTTTTTGTGCAAGAAAAAAGAGACGAGAGGAGAGAGAGAG 194605
QY 301 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 194604 AAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 194564
QY 361 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 194563 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 194504
QY 421 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
Db 194503 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 194444
QY 481 TATTATAGAGAGGTATTTATAGGAGGTATGGGGAATTAAGACAGAGAAACAAATTAG 540
Db 194443 TATTATAGAGAGGTATTTATAGGAGGTATGGGGAATTAAGACAGAGAAACAAATTAG 194384
QY 541 TCCAAAGCAATGATTTCTATTGGGAGTATTTGCCCTAGAGACACTGGCAATACCA 600
Db 194383 TCCAAAGCAATGATTTCTATTGGGAGTATTTGCCCTAGAGACACTGGCAATACCA 194324
QY 601 GGAACATTTTTTGTGTGTCACAATATATAGAGGGGCAATTAAGTATGATAG 660
Db 194323 GGAACATTTTTTGTGTGTCACAATATATAGAGGGGCAATTAAGTATGATAG 194264
QY 661 TGGCAAGTGTGCTTCAACATGCTATGATGACACAGGAGGCTCCACAACCAATTT 720
Db 194263 TGGCAAGTGTGCTTCAACATGCTATGATGACACAGGAGGCTCCACAACCAATTT 194204
QY 721 ATCCACCTTCAGATGCCACAGTGCAGATGCAGAGAACCTTCATCCAGGGCTGAGAAC 780
Db 194203 ATCCACCTTCAGATGCCACAGTGCAGATGCAGAGAACCTTCATCCAGGGCTGAGAAC 194144
QY 781 CGTATTTTTCAGAGAGAGGTATTAAGATGGGTTGTGTGAGAGATGGGAGAGAGGTGT 840
Db 194143 CGTATTTTTCAGAGAGAGGTATTAAGATGGGTTGTGTGAGAGATGGGAGAGAGGTGT 194084
QY 841 GTGTCCAGTAAGAGAAATTAAGGCTTCAGAGCTGAGAGGGAGAGAGTGAAGAGAGAG 900
|||||
```

```
Db 194083 GTGTCCAGTAAGAGAAATTAAGGCTTCAGAGCTGAGAGGGAGAGTGAAGAGAGAG 194024
QY 901 GCGGAGAGATACACATGAGGAGAGACAGCTGGAACAGAAAGTAGAGAGAGAGATTGCA 960
|||||
Db 194023 GCGGAGAGATACACATGAGGAGAGACAGCTGGAACAGAAAGTAGAGAGAGAGATTGCA 193964
QY 961 GATGTGAGAGAGAGAGGTCACAGACCCCGCAAAATGATGTGTGCAACAGAGAAATTGG 1020
|||||
Db 193963 GATGTGAGAGAGAGGTCACAGACCCCGCAAAATGATGTGTGCAACAGAGAAATTGG 193904
QY 1021 AAGAGAAATGAGAGTGGAGAGTGAACAATGGGGTCTTAAGTTGAACCTTGAAGCCAGC 1080
Db 193903 AAGAGAAATGAGAGTGGAGAGTGAACAATGGGGTCTTAAGTTGAAGCCAGC 193844
QY 1081 CATGGTGTCAAGCCTGTAAATCCCAACCTTTGAGAGGTGAGTGGGCAATCACTTGA 1140
|||||
Db 193843 CATGGTGTCAAGCCTGTAAATCCCAACCTTTGAGAGGTGAGTGGGCAATCACTTGA 193784
QY 1141 GGCAGAGATTGAGACACAGCTGGCCAAATGTTGAAACCCGTCCTTACAAAAAAT 1200
Db 193783 GGCAGAGATTGAGACACAGCTGGCCAAATGTTGAAACCCGTCCTTACAAAAAAT 193724
QY 1201 ACAAAAATTTAGCCGGGTGTGGTATGAGACACCTGTAGTCAAGCTACTTGGAGGCTGA 1260
Db 193723 ACAAAAATTTAGCCGGGTGTGGTATGAGACACCTGTAGTCAAGCTACTTGGAGGCTGA 193664
QY 1261 GGCAGAGAAATTTCTTGAACCCCGGAGATGAGAGGCTGCAAGTGAAGTGAAGCTGAGCCACT 1320
Db 193663 GGCAGAGAAATTTCTTGAACCTTGGAGATGAGAGGCTGCAAGTGAAGTGAAGCTGAGCCACT 193604
QY 1321 GCGCTCAACCTGGGCAACAGAGTAGTCCATCTC-AAAAAAAAAAGCTGGATTGG 1379
Db 193603 GCGCTCAACCTGGGCAACAGAGTAGTCCATCTC-AAAAAAAAAAGCTGGATTGG 193544
QY 1380 GAGTGAATATTAATTAATTAATTTCCCTCTCTCCCTTTGGCTGTGCTCCTCTGTC 1439
Db 193543 GAGTGAATATTAATTAATTAATTTCCCTCTCTCTCTCTTTGGCTGTGCTCCTCTGTC 193484
QY 1440 TTTTCTGATTTCTTCACTCTCTGTAATCTTCCATCTCTGTGTGCTTCCATCTGCTT 1499
Db 193483 TTTTCTGATTTCTTCACTCTCTGTAATCTTCCATCTCTGTGTGCTTCCATCTGCTT 193424
QY 1500 CTCATCTATGGGCACTCTGGGTCTCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1559
Db 193423 CTCATCTATGGGCACTCTGGGTCTCTATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 193364
QY 1560 TGCCTCTCATAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1619
Db 193363 TGCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 193304
QY 1620 AATGCTCCCTCAGAGGACTGTGTCTCTGAGAGATTAACCTTGTGCCCGGCCCAACAGA 1679
Db 193303 AATGCTCCCTCAGAGGACTGTGTCTCTGAGAGATTAACCTTGTGCCCGGCCCAACAGA 193244
QY 1680 CCGGGTGTCTACAGCAACTCTGCAAGTTCCACCAAGTGGATCCAGAGAAACATCCAGGCC 1739
Db 193243 CCGGGTGTCTACAGCAACTCTGCAAGTTCCACCAAGTGGATCCAGAGAAACATCCAGGCC 193184
QY 1740 AACTCTGATCATCCAGAGACTCAGACACACCGGATCCCACTCTCTCTCTCTCTCTCTCT 1799
Db 193183 AACTCTGATCATCCAGAGACTCAGACACACCGGATCCCACTCTCTCTCTCTCTCTCTCTCT 193124
QY 1800 CTGACACTCTCTTACAGACCCCTCATTTCTTCCAGAGATTTGAGATGTTCTCTCTCA 1859
Db 193123 CTGACACTCTCTTACAGACCCCTCATTTCTTCCAGAGATTTGAGATGTTCTCTCTCA 193064
QY 1860 GCCCTGACCCCATGCTCTCTGAGACTCAGGGTCTGCTTCCCAATTTGGGCTGACCGTG 1919
Db 193063 GCCCTGACCCCATGCTCTCTGAGACTCAGGGTCTGCTTCCCAATTTGGGCTGACCGTG 193004
QY 1920 TCTCTCTAGTTGAACCTTGGAGAAATTTCCAAAATGTCAGAGGGGGGGGTTGGGTCTC 1979
Db 193003 TCTCTCTAGTTGAACCTTGGAGAAATTTCCAAAATGTCAGAGGGGGGGGTTGGGTCTC 192944
```

```

OY      1980 AATTCCTGGGACATTCATCCAGCTCAGGCGCATCCTCTCGACGCTGA 2039
          |||||||
DB      192943 AATTCCTGGGACATTCATCCAGCTCAGGCGCATCCTCTCGACGCTGA 192884
          |||||||
OY      2040 CCCAATTAGTCCGAGAAATAACTAGAG 2071
          |||||||
DB      192883 CCCAATTAGTCCGAGAAATAACTAGAG 192852
          |||||||

RESULT 5
AC130782/c
LOCUS    200792 bp      DNA      linear      HTG 22-NOV-2002
DEFINITION
Pan troglodytes clone CH251-355A20, WORKING DRAFT SEQUENCE, 12
ordered pieces.
AC130782
AC130782.2 GI:25167101
HTG: HTGS_PHASE2: HTGS_DRAFT.
KEYWORDS
Pan troglodytes (chimpanzee)
SOURCE
Pan troglodytes
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
1 (bases 1 to 200792)
REFERENCE
AUTHORS
Akheri N., Antonellis A., Ayele K., Beckstrom-Sternberg S.M.,
Benjamin B., Blakesley R.W., Bouffard G.G., Brinkley C., Brooks S.,
Carlaigh K., Coleman B., Engle J., Granite S., Guan X., Gupta J.,
Haghighi P., Han J., Hansen N., Ho S.-L., Idol J.R., Karlls E.,
Laric P., Lee-Lin S.-O., Legaspi R., Maduro Q.L., Maduro Y.B.,
Marulles E.H., Mastello C., Maskeri B., McDowell D.,
Paguirigan C., Pearson R., Portnoy M.E., Prasad A.,
Reddy-Dugue N., Schandler K., Schueler M.G., Sison C.,
Starthrop S., Thomas J.W., Thomas P.J., Touchman J.W., Vogt J.L.,
Wetherby K.D., Wiggins L., Young A. and Green E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 200792)
JOURNAL
Direct Submission
Green E.D.
REFERENCE
AUTHORS
Submitted (14-AUG-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 200792)
JOURNAL
Direct Submission
Submitted (22-NOV-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Nov 22, 2002 this sequence version replaced gi:22218452.
COMMENT
----- Genome Center
Center: NIH Intramural Sequencing Center
Vector code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc-zoo@ngri.nih.gov
----- Project Information
Center project name: dhx
Center clone name: 355A20

```

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 19656 bases at least Q40  
Consensus quality: 19783 bases at least Q30

```

FEATURES
source
1.200792
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-355A20"
/clone_1fb="CH251"
1.28306
/notes="assembly_fragment"
clone_end:sp6
vector_side:left"
misc_feature
28407..37856
/notes="assembly_fragment"
37957..73522
/notes="assembly_fragment"
73623..83567
/notes="assembly_fragment"
83668..88817
/notes="assembly_fragment"
88918..125611
/notes="assembly_fragment"
125712..159879
/notes="assembly_fragment"
159980..174698
/notes="assembly_fragment"
174799..186382
/notes="assembly_fragment"
186483..193344
/notes="assembly_fragment"
193445..199363
/notes="assembly_fragment"
199464..200792
/notes="assembly_fragment"
clone_end:t7
vector_side:right"

```

Consensus quality: 198879 bases at least Q20  
Insert size: 215000; agarose-fp  
Insert size: 199692; sum-of-contigs  
Quality coverage: 9.03x in Q20 bases; agarose-fp  
Quality coverage: 9.73x in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
1 28306: contig of 28306 bp in length  
\* 28307 28406: gap of unknown length  
\* 28407 37856: contig of 9450 bp in length  
\* 37857 37956: gap of unknown length  
\* 37957 73522: contig of 35566 bp in length  
\* 73523 73622: gap of unknown length  
\* 73623 83567: contig of 9945 bp in length  
\* 83568 83667: gap of unknown length  
\* 83668 88817: contig of 5150 bp in length  
\* 88818 88917: gap of unknown length  
\* 88918 125611: contig of 36694 bp in length  
\* 125612 125711: gap of unknown length  
\* 125712 159879: contig of 34168 bp in length  
\* 159880 159979: gap of unknown length  
\* 159980 174698: contig of 14719 bp in length  
\* 174699 174798: gap of unknown length  
\* 174799 186382: contig of 11584 bp in length  
\* 186383 186482: gap of unknown length  
\* 186483 193344: contig of 6862 bp in length  
\* 193345 193444: gap of unknown length  
\* 193445 199363: contig of 5919 bp in length  
\* 199364 199463: gap of unknown length  
\* 199464 200792: contig of 1329 bp in length.  
Location/Qualifiers

BASE COUNT	49975 a	49805 c	48502 g	51409 t	1101 others
ORIGIN					
Query Match	81.4%	Score 1686.8	DB 2	Length 200792	
Best Local Similarity	92.0%	Pred. No. 0			
Matches 1907	Conservative	0	Mismatches 37	Indels 128	Gaps 7
QY	1	AAACAGACCAAAACTCTCTGCGCGGTGACCTCATGTGTCCCAAGTGGAGACAGCA	60		
DB	69667	AAACAGACCAAAACTCTCTGCGCGGTGACCTCATGTGTCCCAAGTGGAGACAGCA	69608		
QY	61	TAAAGATAGTAAATATGTAGTAAATTAATAAATAAATAAATAAATAAATAAATAA	120		
DB	69607	TAAAGATAGTAAATATGTAGTAAATTAATAAATAAATAAATAAATAAATAAATAA	69552		
QY	121	CTTGACCTGTAGTCCAGTCTGAGGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG	180		
DB	69551	CTTGACCTGTAGTCCAGTCTGAGGAGGTGAGGTGAGGTGAGGTGAGGTGAGGTGAG	69492		
QY	181	CGTTTGAGGCTGCGGTGAGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	240		
DB	69491	CGTTTGAGGCTGCGGTGAGGAGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	69432		
QY	241	CAAGCAAGACGTTTTTGTGCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA	300		
DB	69431	CAAGCAAGACGTTTTTGTGCAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA	69374		
QY	301	AAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA	360		
DB	69373	-----	69374		
QY	361	GAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA	420		
DB	69373	-----	69371	-AGG	69371
QY	421	GAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA	480		
DB	69370	GAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA	69311		
QY	481	TATTTAGGAGGATATATAGGAGGTATGGGAAATTAAGAAACAAATTAAG	540		
DB	69310	TATTTAGGAGGATATATAGGAGGTATGGGAAATTAAGAAACAAATTAAG	69251		
QY	541	TCCAAGCAATGATTTCTATTTGGAGTGTCTGCCCCCTAGAGACACTGGCAATTA	600		
DB	69250	TCCAAGCAATGATTTCTATTTGGAGTGTCTGCCCCCTAGAGACACTGGCAATTA	69191		
QY	601	GGAGCAATTTTGGTGTACAACTATATGAGGGGCAATTAATGATGATGATGATG	660		
DB	69190	GGAGCAATTTTGGTGTACAACTATATGAGGGGCAATTAATGATGATGATGATG	69131		
QY	661	TGCCAAGTGTCTGTTTCAACATGCTATGATGCACACGAGGCTCCCAACAACTAT	720		
DB	69130	TGCCAAGTGTCTGTTTCAACATGCTATGATGCACACGAGGCTCCCAACAACTAT	69071		
QY	721	ATCCAGCTTCAGATGCCACAGTGCAGATCGAGAAACCTCATCCAGGGCTGAGAC	780		
DB	69070	ATCCAGCTTCAGATGCCACAGTGCAGATCGAGAAACCTCATCCAGGGCTGAGAC	69011		
QY	781	CGTATTTTTCGAAAGGAGGTATTAAGATGGTGTGTGAGATGGGGAAGGAAGGT	840		
DB	69010	CGTATTTTTCGAAAGGAGGTATTAAGATGGTGTGTGAGATGGGGAAGGAAGGT	68951		
QY	841	GTTGCTCAGTAAGAGAAATTAAGGCTTCGACAGGCTGAGGGGAGATGAGGAAGGA	900		
DB	68950	GTTGCTCAGTAAGAGAAATTAAGGCTTCGACAGGCTGAGGGGAGATGAGGAAGGA	68891		
QY	901	GCGGAGAGATACAGATGAGGAGACAGGCTGGAACAAAGTATGAGACGAAGATTGA	960		
DB	68890	GCGGAGAGATACAGATGAGGAGACAGGCTGGAACAAAGTATGAGACGAAGATTGA	68831		
QY	961	GATGTGAGAGAGAGGCTACAGACCCCGGAAATGATGTGTGAGCAACAGAAATCGG	1020		
DB	68830	GATGTGAGAGAGAGGCTACAGACCCCGGAAATGATGTGTGAGCAACAGAAATCGG	68773		
QY	1021	AAGAGGAAGATGAGTGGAGATGACAAATGGGGCTTAAGGTTGAATTTGAGGCCAGG	1080		
DB	68772	AAGAGGAAGATGAGTGGAGATGACAAATGGGGCTTAAGGTTGAATTTGAGGCCAGG	68713		
QY	1081	CATGATGCTCAGCCTGTATATCCCAACACTTTTGGAGGCTGAGGTGGGCAATCAT	1140		
DB	68712	CATGATGCTCAGCCTGTATATCCCAACACTTTTGGAGGCTGAGGTGGGCAATCAT	68653		
QY	1141	GCGCAGAGGTTCGAGACCAAGCTGCGCAACATGCTGTAACCCCGCTCTCTC-AAAA	1199		
DB	68652	GCGCAGAGGTTCGAGACCAAGCTGCGCAACATGCTGTAACCCCGCTCTCTC-AAAA	68593		
QY	1200	TACAAAAATTAAGCGGGTGTGTGATGAGACCTGTAGTACAGCTACTTGGAGGCTG	1259		
DB	68592	TACAAAAATTAAGCGGGTGTGTGATGAGACCTGTAGTACAGCTACTTGGAGGCTG	68533		
QY	1260	AGGCAAGGAATTCCTTGAAACCGGGAGATGAGGCTGCACTGAGTCAAGGCTAC	1319		
DB	68532	AGGCAAGGAATTCCTTGAAACCGGGAGATGAGGCTGCACTGAGTCAAGGCTAC	68473		
QY	1320	TGCGTCCAACTGGGCAACAGATTAAGACTCCATCCAAAAAAGGCTGATTTG	1379		
DB	68472	TGCGTCCAACTGGGCAACAGATTAAGACTCCATCCAAAAAAGGCTGATTTG	68415		
QY	1380	GAGTGAATTAATTAACATTCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1439		
DB	68414	GAGTGAATTAATTAACATTCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	68355		
QY	1440	TTTTTCTCATTTCTCTCATCTCTGACTTTCATCTCTGTGTGTGTGTGTGTGTGT	1499		
DB	68354	TTTTTCTCATTTCTCTCATCTCTGACTTTCATCTCTGTGTGTGTGTGTGTGTGT	68295		
QY	1500	CTCATCTATGAGGATCTCGGGTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1559		
DB	68294	CTCATCTATGAGGATCTCGGGTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT	68235		
QY	1560	TGCCCTCTCATGCGCCCTTTCTCTCTCTGAGGATTTCTGGGGGCTGTGTGTCTG	1619		
DB	68234	TGCCCTCTCTCATGCGCCCTTTCTCTCTCTGAGGATTTCTGGGGGCTGTGTGTCTG	68175		
QY	1620	AATGCTCTCTGACAGGACTGTCTCTGGGAGATTAACCTTGTGCCCCCAACAGA	1679		
DB	68174	AATGCTCTCTGACAGGACTGTCTCTGGGAGATTAACCTTGTGCCCCCAACAGA	68115		
QY	1680	CGGGTCTCTACAGAACCTCTGCAAGTTTCAAGTGATCGAGAAACATCCAGAGCC	1739		
DB	68114	CGGGTCTCTACAGAACCTCTGCAAGTTTCAAGTGATCGAGAAACATCCAGAGCC	68055		
QY	1740	AACTCTGAGTATCCAGAGCTGACAGCACCGGATCCCACTGCTGACAGGACAGCC	1799		
DB	68054	AACTCTGAGTATCCAGAGCTGACAGCACCGGATCCCACTGCTGACAGGACAGCC	67995		
QY	1800	CTGACACTCTTTTACAGACCTCTATTTCTTCCAGAGATGTTGAATAATTTATCTTCA	1859		
DB	67994	CTGACACTCTTTTACAGACCTCTATTTCTTCCAGAGATGTTGAATAATTTATCTTCA	67935		
QY	1860	GCCCCGACCCCATCTCTCTGAGCTGAGGCTGCTTCCCAATTTGGGCTACCCGTG	1919		
DB	67934	GCCCCGACCCCATCTCTCTGAGCTGAGGCTGCTTCCCAATTTGGGCTACCCGTG	67875		
QY	1920	TCTCTCTAGTTGAACCTCGGGAACAATTTCCAAAACTGTCCAGAGCGGGGTTGCTCTC	1979		
DB	67874	TCTCTCTAGTTGAACCTCGGGAACAATTTCCAAAACTGTCCAGAGCGGGGTTGCTCTC	67815		
QY	1980	AATCTCCCTGGGCACTTTCAATCTCAAGCTCAGGGCCATCTCTCTCTCTCTCTCTCA	2039		
DB	67814	AATCTCCCTGGGCACTTTCAATCTCAAGCTCAGGGCCATCTCTCTCTCTCTCTCTCA	67755		
QY	2040	CCCAATTTTATGCCAGAAATTAACGTAGAG	2071		
DB	67754	CCCAATTTTATGCCAGAAATTAACGTAGAG	67723		



REFERENCE 1 (bases 1 to 586)  
AUTHORS Myers R.M.  
TITLE Human STS (1998)  
JOURNAL Unpublished (1998)  
COMMENT

Contact: Richard M. Myers  
Stanford Human Genome Center (SHGC)  
Stanford University School of Medicine  
Department of Genetics, M-344, Stanford, CA 94305, USA  
Tel: 4157259687  
Fax: 4157259689  
Email: myers@shgc.stanford.edu  
Primer A: AGAGACAGCGTCAGCCCAAT  
Primer B: GCCAATCTGTGATCATCCC  
STS size: 188  
PCR Profile:

Initial incubation: 95 degrees C for 10 minutes

## Denaturation:

94 degrees C for 30 seconds  
60 degrees C for 30 seconds

## Annealing:

72 degrees C for 23 seconds

## Polymerization:

30

## PCR Cycles:

Perkin Elmer 9700

## Template:

25 ng

## Primer:

each 1 uM

## dNTPs:

each 200 uM

## Amplifrag Gold Polymerase:

0.07 units/ul

## Total Vol:

5 ul

## Buffer:

MgCl2: 2.5 mM

KCl: 50 mM

Tris-HCl: 10 mM

pH: 8.3

## Prepared with primer pairs derived from W73140 -- Unigene.

## Location/Qualifiers

1..586

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/map="19"

/clone\_idb="Human"

STS

primer\_bind

primer\_bind

BASE COUNT

Query Match 23.3%: Score 482; DB 11; Length 586;  
Best Local Similarity 100.0%; Pred. No. 8.1e-111; Mismatches 0; Indels 0; Gaps 0;  
Matches 482; Conservative 0;

1590 CAGGGTATTCGCGGGGCTGTGTCGTCATGTCCTCCCTCAGGAGCTGTCCTGG 1649  
|||||  
501 CAGGGGATTCGCGGGGCTGTGTCGTCATGTCCTCCCTCAGGAGCTGTCCTGG 442  
|||||  
1650 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTTGCAGTTC 1709  
|||||  
441 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTTGCAGTTC 382  
|||||  
1710 ACCAAGTGTGATCCAGGAACATCCAGGCACTCTGATCCAGGAGTCAAGCA 1769  
|||||  
381 ACCAAGTGTGATCCAGGAACATCCAGGCACTCTGATCCAGGAGTCAAGCA 322  
|||||  
1770 CCGGATCCCACTGCTGTCAGAGGAGAGCCCTGACACTCTTTCAGAGCCCTGATTCCTTC 1829  
|||||  
321 CCGGATCCCACTGCTGTCAGAGGAGAGCCCTGACACTCTTTCAGAGCCCTGATTCCTTC 262  
|||||  
1830 CCAGAGATGTTGAGATGTTCTATCTCTCCAGCCCTGACCCCATGTCCTGAGCTCAGG 1889  
|||||

Db 261 CCAGAGATGTTGAGATGTTCTATCTCTCCAGCCCTGACCCCATGTCCTGAGCTCAGG 202  
QY 1890 GTCGCTCCCACTGCTGTCAGAGGAGAGCCCTGACACTCTTTCAGAGCCCTGATTCCTTC 1949  
|||||  
Db 201 GTCGCTCCCACTGCTGTCAGAGGAGAGCCCTGACACTCTTTCAGAGCCCTGATTCCTTC 142  
|||||

QY 1950 CAAACTGTCCAGAGGCGGGGCTGTGTCGTCATCTCCCTGAGGAGCTTTCATCTCAAGC 2009  
|||||

Db 141 CAAACTGTCCAGAGGCGGGGCTGTGTCGTCATCTCCCTGAGGAGCTTTCATCTCAAGC 82  
|||||

QY 2010 TCAGGCGCCATCCCTTCTCTGTCAGCTGTCAGCCCAATTTATCCCAAGAAATTAAGTGA 2069  
|||||

Db 81 TCAGGCGCCATCCCTTCTCTGTCAGCTGTCAGCCCAATTTATCCCAAGAAATTAAGTGA 22  
|||||

QY 2070 AG 2071  
||

Db 21 AG 20

RESULT 9

AX429955/c 735 bp DNA Linear PAT 21-JUN-2002

LOCUS AX429955 Sequence 94 from Patent WO0198339.

DEFINITION AX429955

ACCESSION AX429955

VERSION AX429955.1 GI:21541119

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

1

Reed,S.G., Xu,J., Dillon,D.C., Retter,M.W. and Harlocker,S.L.

Compositions and methods for the therapy and diagnosis of breast

Cancer

Patent: WO 0198339-A 94 27-DEC-2001;

CORTIXA CORPORATION (US)

Location/Qualifiers

1..735

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT

161 a 176 c 226 g 172 t

ORIGIN

Query Match 23.3%: Score 482; DB 6; Length 735;

Best Local Similarity 100.0%; Pred. No. 8.3e-111; Mismatches 0; Indels 0; Gaps 0;

Matches 482; Conservative 0;

1590 CAGGGTATTCGCGGGGCTGTGTCGTCATGTCCTCCCTCAGGAGCTGTCCTGG 1649  
|||||  
500 CAGGGTATTCGCGGGGCTGTGTCGTCATGTCCTCCCTCAGGAGCTGTCCTGG 441  
|||||  
1650 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTTGCAGTTC 1709  
|||||  
440 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTTGCAGTTC 381  
|||||  
1710 ACCAAGTGTGATCCAGGAACATCCAGGCACTCTGATCCAGGAGTCAAGCA 1769  
|||||  
380 ACCAAGTGTGATCCAGGAACATCCAGGCACTCTGATCCAGGAGTCAAGCA 321  
|||||  
1770 CCGGATCCCACTGCTGTCAGAGGAGAGCCCTGACACTCTTTCAGAGCCCTGATTCCTTC 1829  
|||||  
320 CCGGATCCCACTGCTGTCAGAGGAGAGCCCTGACACTCTTTCAGAGCCCTGATTCCTTC 261  
|||||  
1830 CCAGAGATGTTGAGATGTTCTATCTCTCCAGCCCTGACCCCATGTCCTGAGCTCAGG 1889  
|||||  
260 CCAGAGATGTTGAGATGTTCTATCTCTCCAGCCCTGACCCCATGTCCTGAGCTCAGG 201  
|||||  
1890 GTCGCTCCCACTGCTGTCAGAGGAGAGCCCTGACACTCTTTCAGAGCCCTGATTCCTTC 1949  
|||||  
200 GTCGCTCCCACTGCTGTCAGAGGAGAGCCCTGACACTCTTTCAGAGCCCTGATTCCTTC 141  
|||||  
1950 CAAACTGTCCAGAGGCGGGGCTGTGTCGTCATCTCCCTGAGGAGCTTTCATCTCAAGC 2009  
|||||

|||||  
Db 140 CAAAGTCGACAGGGGGGGGGTCCGCTCAATCTCCCTGGGACACTTTCATCTCAAGC 81  
QY 2010 TCAGGGCCCATCCCTCTCTGTCAGCTGTGACCCAAATTTAGTCCAGAAATAACTGAGA 2069  
Db 80 TCAGGGCCCATCCCTCTCTGTCAGCTGTGACCCAAATTTAGTCCAGAAATAACTGAGA 21  
QY 2070 AG 2071  
||  
Db 20 AG 19  
RESULT 10  
BD139877/c 738 bp DNA linear PAT 18-SEP-2002  
LOCUS Compounds for immunotherapy and diagnosis of breast cancer and  
DEFINITION methods for their use.  
ACCESSION BD139877  
VERSION BD139877.1 GI:23234822  
KEYWORDS JP 2002507387-A/94.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 738)  
AUTHORS Reed,S.G. and Xu,J.  
TITLE Compounds for immunotherapy and diagnosis of breast cancer and  
JOURNAL methods for their use  
PATENT: JP 2002507387-A 94 12-MAR-2002;  
CORIXA CORP  
OS Homo sapiens (human)  
PN JP 2002507387-A/94  
PD 12-MAR-2002  
PR 22-DEC-1998 JP 2000526543  
PR 24-DEC-1997 US 08/998253,24-DEC-1997 US 08/998255 PR  
17-JUL-1998 US 09/118627,17-JUL-1998 US 09/118554 PI STEVEN  
G REED,/JIANGHUN XU  
PC C12N15/09,A61K38/00,A61K39/39,A61K39/395,A61P35/00,  
PC C07K14/47,  
PC C07K16/18,C12N1/19,C12N1/21,C12N5/10,C12P21/08,C12Q1/68,G01N33/PC  
53,  
PC G01N33/577,C12N15/00,A61K37/02,C12N5/00  
CC Compounds for immunotherapy and diagnosis  
of breast cancer and  
CC their use methods for  
CC key Location/Qualifiers  
FH source 1..738  
FT Location/Qualifiers  
1..738  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
BASE COUNT 161 a 179 c 226 g 172 t  
ORIGIN  
Query Match 23.3%; Score 482; DB 6; Length 738;  
Best Local Similarity 100.0%; Pred. No. 8.3e-111;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1770 CGGGCAATCCCACTGCTGACGGAGACGCCCTGACACTCTCTTCAGACCTCATTCCTTC 1829  
Db 323 CGGGCAATCCCACTGCTGACGGAGACGCCCTGACACTCTCTTCAGACCTCATTCCTTC 264  
QY 1830 CCAGAGATGTTGGAATGTTCAATCTCTCCAGCCCTGACCCCATATGTCCTCGAGCTCAG 1889  
Db 263 CCAGAGATGTTGGAATGTTCAATCTCTCCAGCCCTGACCCCATATGTCCTCGAGCTCAG 204  
QY 1890 GTCTGCTTCCCAATGTTGGCTGACCGTGTCTCTAGTTGAACCTGGGAAATTC 1949  
Db 203 GTCTGCTTCCCAATGTTGGCTGACCGTGTCTCTAGTTGAACCTGGGAAATTC 144  
QY 1950 CAAAGTCGACAGGGGGGGGGTCCGCTCAATCTCCCTGGGACACTTTCATCTCAAGC 2009  
Db 143 CAAAGTCGACAGGGGGGGGGTCCGCTCAATCTCCCTGGGACACTTTCATCTCAAGC 84  
QY 2010 TCAGGGCCCATCCCTCTCTGTCAGCTGTGACCCAAATTTAGTCCAGAAATAACTGAGA 2069  
Db 83 TCAGGGCCCATCCCTCTCTGTCAGCTGTGACCCAAATTTAGTCCAGAAATAACTGAGA 24  
QY 2070 AG 2071  
||  
Db 23 AG 22  
RESULT 11  
AY279380  
LOCUS 1370 bp mRNA linear PRI 26-MAY-2003  
DEFINITION Homo sapiens kallikrein 5 splice variant 1 (KLK5) mRNA, complete  
ACCESSION AY279380  
VERSION AY279380.1 GI:31075480  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1370)  
AUTHORS Kurliender,L., Yousef,G.M., White,N.M.A., Robb,J.-D., Borgono,C.A.  
and Diamandis,E.P.  
TITLE Identification of splice variants for the human kallikrein gene 5  
JOURNAL (KLK5)  
REFERENCE 2 (bases 1 to 1370)  
AUTHORS Kurliender,L., Yousef,G.M., White,N.M.A., Robb,J.-D., Borgono,C.A.  
and Diamandis,E.P.  
TITLE Direct Submission  
JOURNAL Submitted (19-APR-2003) Pathology and Laboratory Medicine, Mount  
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,  
Canada  
FEATURES  
source Location/Qualifiers  
1..1370  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13.4"  
1..1370  
/gene="KLK5"  
1..127  
/gene="KLK5"  
128..1009  
/gene="KLK5"  
/note="serine protease; hK5; alternatively spliced;  
similar to the product of Genbank Accession Number  
AF135028"  
/codon\_start=1  
/product="kallikrein 5 splice variant 1"  
/protein\_id="AAP42275.1"  
/db\_xref="GI:31075481"  
/translation="MATARPMMVLCALITALLGVTEHVLANNVSCDHSNTPVS  
GSMQDVGAGGEDARSDSSRLINSDDMHTQPMQALLLRPNQLYCGAVLVHWPQ

LTAAHCKRRKFRVRLGHVSLSPVYESQOMFOGVKSIPHPGYSHPGSHNDMLIKLN  
 DAIPTKDVRLPINSVSHKPSAGTCKLVSWMGTGTSKPOVHPFKVLOCCLNITSVLSORCE  
 DAYPRIDDTMTEFCAGKAGRDSQSDSGSPVVCNSLGLVSWBDYPCARRNRPQVYT  
 NICKRTKNIQETIQANS"

3'UTR  
 polyA-signal  
 /gene="RLK5"  
 1010..1370  
 1319..1324  
 /gene="RLK5"

BASE COUNT 323 a 424 c 352 g 271 t  
 ORIGIN

Query Match 23.3%; Score 482; DB 9; Length 1370;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-111;  
 Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGTGATTCGGGGGGGCGCTGTGCTGCATGGCTCCCTGCAGGAGCTGCTGCTGG 1649  
 DB 851 CAGGGTGATTCGGGGGGGCGCTGTGCTGCATGGCTCCCTGCAGGAGCTGCTGCTGG 910  
 QY 1650 GGAGATTACCTTGTGCCGGGCCCAACAGACCGGGGTGTCTACAGACCTCTGCAAGTTTC 1709  
 DB 911 GGAGATTACCTTGTGCCGGGCCCAACAGACCGGGGTGTCTACAGACCTCTGCAAGTTTC 970  
 QY 1710 ACCAAGTGGATCCAGAAACCATCCAGGCCAATCCTGAGTCATGCCAGAGACTGACACA 1769  
 DB 971 ACCAAGTGGATCCAGAAACCATCCAGGCCAATCCTGAGTCATGCCAGAGACTGACACA 1030  
 QY 1770 CCGGATCCCGACCTGCTGCAGAGGACAGCCCTGACACCTCTTTCAGACCTCTATCTTC 1829  
 DB 1031 CCGGATCCCGACCTGCTGCAGAGGACAGCCCTGACACCTCTTTCAGACCTCTATCTTC 1090  
 QY 1830 CCAGAGATTGAGATGTTATCTCTCCAGGCCCTGACCCCATGCTCTCTGAGTCTGAGTCAAG 1889  
 DB 1091 CCAGAGATTGAGATGTTATCTCTCCAGGCCCTGACCCCATGCTCTCTGAGTCTGAGTCAAG 1150  
 QY 1890 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCTGGGAACAATTTC 1949  
 DB 1151 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCTGGGAACAATTTC 1210  
 QY 1950 CAAAAGTGTCCAGAGGGGGGGGTGGTGTCTCATCTCCCGGGGGCACTTTCATCTCAAGC 2009  
 DB 1211 CAAAAGTGTCCAGAGGGGGGGGTGGTGTCTCATCTCCCGGGGGCACTTTCATCTCAAGC 1270  
 QY 2010 TCAGGGCCCATCTCTCTCTGAGCTGTGACCCCAATTAGTCCAGAAATAAAGTGAAGA 2069  
 DB 1271 TCAGGGCCCATCTCTCTCTGAGCTGTGACCCCAATTAGTCCAGAAATAAAGTGAAGA 1330  
 QY 2070 AG 2071  
 DB 1331 AG 1332

## RESULT 12.

BD107879 1381 bp DNA linear PAT 18-SEP-2002  
 LOCUS BD107879 36 human secreted proteins.  
 DEFINITION BD107879  
 ACCESSION BD107879.1 GI:23202697  
 VERSION JP 2002500035-A/50.  
 KEYWORDS JP 2002500035-A/50.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1381)

Ruben,S.M., Soppet,D.R., Ebner,R., Lafleur,D.W., Ni,J.,  
 Brewer,L.A., Olsen,H.S., Duan,R.D. and Rosen,C.A.

36 human secreted proteins  
 Patent: JP 2002500035-A 50 08-JAN-2002;

HUMAN GENOME SCIENCES INC

OS Homo sapiens (human)  
 PN JP 2002500035-A/50  
 PD 08-JAN-2002  
 PF 06-JAN-1999 JP 2000527554

COMMENT

PR 07-JAN-1998 US 60/070657, 07-JAN-1998 US 60/070658 PR  
 07-JAN-1998 US 60/070692, 07-JAN-1998 US 60/070704 PI STEVEN  
 M RUBEN, DANIEL R SOPPET, REINHARD EBNER, DAVID W LAFLEUR, PI JIAN  
 NI, LAURIE A BREMER, HENRIK S OLSEN, ROSANNE D DUAN, CRAIG A ROSEN PC  
 PI C12N15/09, A61K31/711, A61K38/00, A61K39/393, A61K39/395, A61K48/00,  
 PC A61P7/00, A61P11/06, A61P19/02, A61P29/00, A61P31/18, A61P35/00, PC  
 A61P35/02,  
 PC A61P37/00, C07K14/435, C07K16/18, C12N1/15, C12N1/21, C12N5/10, PC  
 C12N15/00,  
 PC A61K37/02, C12N5/00  
 CC 36 human secreted proteins  
 FH Key Location/Qualifiers  
 FT source 1..1381  
 FT /organism="Homo sapiens (human)".  
 FEATURES  
 source Location/Qualifiers  
 1..1381  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"

Query Match 23.3%; Score 482; DB 6; Length 1381;  
 Best Local Similarity 100.0%; Pred. No. 8.9e-111;  
 Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGTGATTCGGGGGGGCGCTGTGCTGCATGGCTCCCTGCAGGAGCTGCTGCTGG 1649  
 DB 852 CAGGGTGATTCGGGGGGGCGCTGTGCTGCATGGCTCCCTGCAGGAGCTGCTGCTGG 911  
 QY 1650 GGAGATTACCTTGTGCCGGGCCCAACAGACCGGGGTGTCTACAGACCTCTGCAAGTTTC 1709  
 DB 912 GGAGATTACCTTGTGCCGGGCCCAACAGACCGGGGTGTCTACAGACCTCTGCAAGTTTC 971  
 QY 1710 ACCAAGTGGATCCAGAAACCATCCAGGCCAATCCTGAGTCATGCCAGAGACTGACACA 1769  
 DB 972 ACCAAGTGGATCCAGAAACCATCCAGGCCAATCCTGAGTCATGCCAGAGACTGACACA 1031  
 QY 1770 CCGGATCCCGACCTGCTGCAGAGGACAGCCCTGACACCTCTTTCAGACCTCTATCTTC 1829  
 DB 1032 CCGGATCCCGACCTGCTGCAGAGGACAGCCCTGACACCTCTTTCAGACCTCTATCTTC 1091  
 QY 1830 CCAGAGATTGAGATGTTATCTCTCCAGGCCCTGACCCCATGCTCTCTGAGTCTGAGTCAAG 1889  
 DB 1092 CCAGAGATTGAGATGTTATCTCTCCAGGCCCTGACCCCATGCTCTCTGAGTCTGAGTCAAG 1151  
 QY 1890 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCTGGGAACAATTTC 1949  
 DB 1152 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCTGGGAACAATTTC 1211  
 QY 1950 CAAAAGTGTCCAGAGGGGGGGGTGGTGTCTCATCTCCCGGGGGCACTTTCATCTCAAGC 2009  
 DB 1212 CAAAAGTGTCCAGAGGGGGGGGTGGTGTCTCATCTCCCGGGGGCACTTTCATCTCAAGC 1271  
 QY 2010 TCAGGGCCCATCTCTCTCTGAGCTGTGACCCCAATTAGTCCAGAAATAAAGTGAAGA 2069  
 DB 1272 TCAGGGCCCATCTCTCTCTGAGCTGTGACCCCAATTAGTCCAGAAATAAAGTGAAGA 1331  
 QY 2070 AG 2071  
 DB 1332 AG 1333

## RESULT 13

AF168768 1387 bp mRNA linear PRI 16-APR-2002  
 LOCUS AF168768  
 DEFINITION Homo sapiens stratum corneum tryptic enzyme (SCTE) mRNA, complete  
 cds  
 ACCESSION AF168768  
 VERSION AF168768.2 GI:20153423  
 KEYWORDS



polyA-signal 1387..1392  
/gene="KLK5"  
BASE COUNT 332 a 443 c 374 g 289 t  
ORIGIN

Query Match 23.3%; Score 482; DB 9; Length 1438;  
Best Local Similarity 100.0%; Pred. No. 9e-111;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1590 CAGGTGATTTGGGGGGCCCTGTGCTGCAATGGCTCCCTGCAGGAGACTGTCTCTGG 1649
DB 919 CAGGTGATTTGGGGGGCCCTGTGCTGCAATGGCTCCCTGCAGGAGACTGTCTCTGG 978
QY 1650 GGAGATTACCTTGTGCGCCGCGCCACAGACAGCCGGGTGTCTACAGAACTCTGCAAGTTC 1709
DB 979 GGAGATTACCTTGTGCGCCGCGCCACAGACAGCCGGGTGTCTACAGAACTCTGCAAGTTC 1038
QY 1710 ACCAAGTGTGATCCAGAAACCATCCAGGCCAATCTCTGAGTCAATCCAGAGACTGACACA 1769
DB 1039 ACCAAGTGTGATCCAGAAACCATCCAGGCCAATCTCTGAGTCAATCCAGAGACTGACACA 1098
QY 1770 CCGGATCCCATCCCTGCTGCGAGGAGACAGCCCTGACACTCTTTCAGACCTTCATCTCTTC 1829
DB 1099 CCGGATCCCATCCCTGCTGCGAGGAGACAGCCCTGACACTCTTTCAGACCTTCATCTCTTC 1158
QY 1830 CCAGAGATGTTGAGAAATGTTATCTCTCCAGCCCTGACCCCATGTCCTCTGAGACTCAGG 1889
DB 1159 CCAGAGATGTTGAGAAATGTTATCTCTCCAGCCCTGACCCCATGTCCTCTGAGACTCAGG 1218
QY 1880 GTCTGCTTCCCTCCACATTTGGCTGACCGTGTCTCTAGTTGAACCTCTGGAGCAATTTTC 1949
DB 1219 GTCTGCTTCCCTCCACATTTGGCTGACCGTGTCTCTAGTTGAACCTCTGGAGCAATTTTC 1278
QY 1950 CAAAGCTGTCCAGGGCGGGGGGTGGTGTCAATCTCCCTGGGGGCACTTTCATCTCTCAAGC 2009
DB 1279 CAAAGCTGTCCAGGGCGGGGGGTGGTGTCAATCTCCCTGGGGGCACTTTCATCTCTCAAGC 1338
QY 2010 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTAGTCCAGAAATTAACCTGAGA 2069
DB 1339 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTAGTCCAGAAATTAACCTGAGA 1398
QY 2070 AG 2071
DB 1399 AG 1400
```

RESULT 15  
BD107880 1439 bp DNA linear PAT 18-SEP-2002  
LOCUS 36 human secreted proteins.  
DEFINITION BD107880  
ACCESSION BD107880.1 GI:23202698  
VERSION JP 2002500035-A/51.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 1439)  
AUTHORS Ruben,S.M., Soppet,D.R., Ebner,R., Lafleur,D.W., NI,J.,  
Breuer,L.A., Olsen,H.S., Duan,R.D. and Rosen,C.A.  
36 human secreted proteins  
Patent: JP 2002500035-A 51 08-JAN-2002;  
TITLE HUMAN GENOME SCIENCES INC  
JOURNAL OS Homo sapiens (human)  
COMMENT PN JP 2002500035-A/51  
PD 08-JAN-2002  
PE 06-JAN-1999 JP 2000527554  
PR 07-JAN-1998 US 60/070657, 07-JAN-1998 US 60/070658 PR  
M RUBEN,DANIEL R SOPPET,REINHARD EBNER,DAVID W LAFLEUR, PI JIAN  
NI LAURIE A BREWER,HENRIK S OLSEN,ROSANNE D DUAN,CRAIG A ROSEN PC  
PI C12N15/09,A6IK31/711,A6IK38/00,A6IK39/395,A6IK48/00,

PC A6IP5/00  
PC A6IP7/00,A6IP11/06,A6IP19/02,A6IP29/00,A6IP31/18,A6IP35/00, PC  
A6IP35/02,  
PC A6IP37/00,C07K14/435,C07K16/18,C12N1/15,C12N1/21,C12N5/10, PC  
C12N15/00,  
PC A6IK37/02,C12N5/00  
CC 36 human secreted proteins  
FH key Location/Qualifiers  
FT source 1..1439  
FT 1..1439  
/organism="Homo sapiens (human)".  
FEATURES  
source  
1..1439  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"

BASE COUNT 325 a 446 c 355 g 313 t  
ORIGIN

Query Match 23.3%; Score 482; DB 6; Length 1439;  
Best Local Similarity 100.0%; Pred. No. 9e-111;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1590 CAGGTGATTTGGGGGGCCCTGTGCTGCAATGGCTCCCTGCAGGAGACTGTCTCTGG 1649
DB 910 CAGGTGATTTGGGGGGCCCTGTGCTGCAATGGCTCCCTGCAGGAGACTGTCTCTGG 969
QY 1650 GGAGATTACCTTGTGCGCCGCGCCACAGACAGCCGGGTGTCTACAGAACTCTGCAAGTTC 1709
DB 970 GGAGATTACCTTGTGCGCCGCGCCACAGACAGCCGGGTGTCTACAGAACTCTGCAAGTTC 1029
QY 1710 ACCAAGTGTGATCCAGAAACCATCCAGGCCAATCTCTGAGTCAATCCAGAGACTGACACA 1769
DB 1030 ACCAAGTGTGATCCAGAAACCATCCAGGCCAATCTCTGAGTCAATCCAGAGACTGACACA 1089
QY 1770 CCGGATCCCATCCCTGCTGCGAGGAGACAGCCCTGACACTCTTTCAGACCTTCATCTCTTC 1829
DB 1090 CCGGATCCCATCCCTGCTGCGAGGAGACAGCCCTGACACTCTTTCAGACCTTCATCTCTTC 1149
QY 1830 CCAGAGATGTTGAGAAATGTTATCTCTCCAGCCCTGACCCCATGTCCTCTGAGACTCAGG 1889
DB 1150 CCAGAGATGTTGAGAAATGTTATCTCTCCAGCCCTGACCCCATGTCCTCTGAGACTCAGG 1209
QY 1880 GTCTGCTTCCCTCCACATTTGGCTGACCGTGTCTCTAGTTGAACCTCTGGAGCAATTTTC 1949
DB 1210 GTCTGCTTCCCTCCACATTTGGCTGACCGTGTCTCTAGTTGAACCTCTGGAGCAATTTTC 1269
QY 1950 CAAAGCTGTCCAGGGCGGGGGGTGGTGTCAATCTCCCTGGGGGCACTTTCATCTCTCAAGC 2009
DB 1270 CAAAGCTGTCCAGGGCGGGGGGTGGTGTCAATCTCCCTGGGGGCACTTTCATCTCTCAAGC 1329
QY 2010 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTAGTCCAGAAATTAACCTGAGA 2069
DB 1330 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAATTTAGTCCAGAAATTAACCTGAGA 1389
QY 2070 AG 2071
DB 1390 AG 1391
```

Search completed: August 5, 2003, 09:57:09  
Job time : 7380.88 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 5, 2003, 05:12:55 ; Search time 129.698 seconds  
(without alignments)  
6806.329 Million cell updates/sec

Title: US-09-936-271B-13\_COPY\_1\_2000

Perfect score: 2000

Sequence: 1 gggccagagtgaggcaag.....gtgcgtctgcacccacatc 2000

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB-seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	444	22.2	246240	2	US-08-724-394A-20
C 2	444	22.2	246240	2	US-08-724-394A-21
C 3	444	22.2	246240	2	US-08-724-394A-22
C 4	380.4	19.0	246240	2	US-08-724-394A-20
C 5	380.4	19.0	246240	2	US-08-724-394A-21
C 6	380.4	19.0	246240	2	US-08-724-394A-22
C 7	366.4	18.3	70000	4	US-09-851-896-3
C 8	363.6	18.2	84495	4	US-09-797-906-3
C 9	362.8	18.1	14581	4	US-08-520-373D-4
C 10	362.8	18.1	22481	4	US-08-367-841A-43
C 11	362.8	18.1	22481	5	PCT-US95-07201-43
C 12	362.8	18.1	22484	4	US-09-875-223-2
C 13	333.6	17.7	5375	3	US-08-757-223-7
C 14	333.6	17.7	116592	4	US-09-820-924-3
C 15	333.6	17.2	116592	4	US-09-818-512-3
C 16	339.6	17.0	3805	4	US-09-108-006C-3
C 17	339.2	17.0	168575	4	US-09-426-290-1
C 18	337.2	16.9	29629	4	US-09-729-995-3
C 19	336.6	16.8	43950	4	US-09-735-934A-3
C 20	336.6	16.8	43950	4	US-10-060-332-3
C 21	333.6	16.7	174493	4	US-09-804-471A-3
C 22	332.8	16.6	18073	3	US-09-078-294-12
C 23	332.6	16.6	43950	4	US-09-735-934A-3
C 24	332.6	16.6	43950	4	US-10-060-332-3
C 25	331	16.6	19011	1	US-08-310-356-36
C 26	331	16.6	19557	5	PCT-US92-06300-1
C 27	330.8	16.5	70000	4	US-09-851-896-3

Sequence 10, Appl  
Sequence 15, Appl  
Sequence 18, Appl  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 7, Appl  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 3, Appl  
Sequence 3, Appl  
Sequence 10, Appl  
Sequence 3, Appl  
Sequence 3, Appl  
Sequence 18, Appl  
Sequence 4, Appl  
Sequence 43, Appl  
Sequence 43, Appl  
Sequence 2, Appl

#### ALIGNMENTS

#### RESULT 1

US-08-724-394A-20/c  
; Sequence 20, Application US/08724394A  
; Patent No. 5872237  
; GENERAL INFORMATION:  
; APPLICANT: Feder, John N.  
; APPLICANT: Kronmal, Gregory S.  
; APPLICANT: Lauer, Peter M.  
; APPLICANT: Ruddy, David A.  
; APPLICANT: Thomas, Winston  
; APPLICANT: Tsuchihashi, Zenta  
; APPLICANT: Wolff, Roger K.  
; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
; STREET: Two Embarcadero Center, 8th Floor  
; CITY: San Francisco  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/724,394A  
; FILING DATE: 01-OCT-1996  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Fitts, Renee A.  
; REGISTRATION NUMBER: 35,136  
; REFERENCE/DOCKET NUMBER: 017957-000100  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-576-0200  
; TELEFAX: 415-576-0300  
; INFORMATION FOR SEQ ID NO: 20:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 246240 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: cdna  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 1..246240  
; OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-20











Qy	479	TCCCAAGTAGCTGGGACTACAGGCGCCGCCACTACGCCCGCTAAATTTTTTTGTATTTT	538
Db	13990	TCCCGAGTAGCTGGGACTACAGGCGCCACCACCAATGCTGGTAA-TTTTTTGTATTTT	13932
Qy	539	TACTAGAGAGGGGTTTACCGCTTTAGCCGGGATGGCTCGATCTCTGACCTCGACCTCGTGAT	598
Db	13931	TAGTAGAGAGGGGTTTACCGGTGTTAGCAGGATGGTCTGATCTCTGACCTTGTGAT	13872
Qy	599	CCGCCGCGCTCGCGCTCCCAAAAGTGTGGGATTTACAGGCGTGAGCCACCGCGCCCGGCCA	658
Db	13871	CCACCGCTCGGTCTACAAAGTCTGGGATTTACAGGCATGAGCCACCGCGCCACG--	13814
Qy	659	TGATCATCTCTTGACTATGCTGTGATGTGACAAAGTACCTAAAGCCATCAGACTCTACCCCT	718
Db	13813	-----CCGATGACCCACTTTTTTAAAAACCATCA-----	13786
Qy	719	TAAATATGCAGTTTGGCGCAGCACCGTGGCTCATGCTGTAAATTCACGACTTTGGGAG	778
Db	13785	-----GCCGAGTGGCTACACCTTATATCCACGACTTTGGGAG	13747
Qy	779	GCAGAGGTGGTGAATCACTTTGAGCCAGGAGTTTTGAGACCAAGCTGCCCAACATGTTGA	838
Db	13746	GCTGAGGTGGTGAATCACTTTGAGGTTCAGGAGTTTCAGACCATCTTGGCAACACGGTGA	13687
Qy	839	AACCTCTGCTTTACTAAAAAATAAAAAAATAAAAAAATACAGCCGGGTGTCTGGGGC	898
Db	13686	AAACCCGCTCTCT-----ACTAAAAATACAAAAAATTTAAACGGGGCATTTGGTGTGC	13638
Qy	899	ACACCTGTATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTTGAACCTTGGAGG	958
Db	13637	ATGCCCTGTATCCAGCTACTCAGGAGGCTGAGGCACGAGATTCGCTTTGAACCCAGGAGG	13578
Qy	959	CGGAGGTTCAGTGGCGCCGAGATCACATCACCGCCCTCCAGCCTGGGGCGACAGCAGAAGA	1018
Db	13577	CAGAGGTTCAGTGAACCTGAGATAGTGCACCTGTACTCCAGCCTGGTGACAGAGTGAGA	13518
Qy	1019	CTCTGTCTCAATATAATAATTAACAACAGCA	1050
Db	13517	CCCTGTCTCAAAAAATAAATAAGTAAATAAAA	13486

RESULT 10  
IIS-08-367-841A-43/C

US-08-367-841A-43/C  
 : Sequence 43, Application US/08367841A  
 : Patent No. 6319687  
 :  
 : GENERAL INFORMATION:  
 :  
 : APPLICANT: Chader, Gerald J.; Rodd  
 : APPLICANT: Ignacio R.; Mazuruk, K  
 : APPLICANT: Tombran-Tink, Joyce  
 : TITLE OF INVENTION: PIGMENT EPIPH  
 : TITLE OF INVENTION: DERIVED FACTO  
 : TITLE OF INVENTION: ORGANIZATION  
 : NUMBER OF SEQUENCES: 43  
 : CORRESPONDENCE ADDRESS:  
 : ADDRESSEE: Morgan & Finnegan  
 : STREET: 345 Park Avenue  
 : CITY: New York  
 : STATE: New York  
 : COUNTRY: USA  
 : ZIP: 10154  
 :  
 : COMPUTER READABLE FORM:  
 : MEDIUM TYPE: Floppy Disk  
 : COMPUTER: IBM PC Compatible  
 : OPERATING SYSTEM: PC-DOS/MS-DOS  
 : SOFTWARE: WORDPERFECT 5.1  
 :  
 : CURRENT APPLICATION DATA:  
 : APPLICATION NUMBER: US/08/367.8  
 : FILING DATE: 30-DEC-1994  
 : CLASSIFICATION: 435  
 :  
 : PRIOR APPLICATION DATA:  
 : APPLICATION NUMBER: 08/257,963  
 : FILING DATE: 07-JUN-1994

Query Match	18.1%;	Score 362.8;	DB 4;	Length 14581;
Best Local Similarity	76.3%;	Pred. No. 5.4e-65;		
Matches 528;	Conservative 0;	Mismatches 97;	Indels 67;	Gaps 4;

  

Qy	359	TTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTGCCCCAGGCTCGAGATGCAGTGGC	418
Db	14110	TTTTTTTTTTTTTTTTTTGAGACGGAGTCTCGCTCTGTGCTCAGGCTCGAGTACGGTGGC	14051

  

Qy	419	GGGATCTCGGCTACTGCAAGCTCGCCCTCCGGGTTCAAGCCATTCTTCGCTTCAGCC	478
Db	14050	GTGATCTCGGCTCACTGCAAGCTCGCCCTCCAGGTTCAAGCCATTCTTCCTCGCTTCAGCC	13991





Db 1767 GTGTGGTGTGC 1803

RESULT 14  
US-09-820-924-3  
; Sequence 3, Application US/09820924  
; Patent No.. 6555351  
; GENERAL INFORMATION:  
; APPLICANT: BEASLEY, Ellen M. et al  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001213  
; CURRENT APPLICATION NUMBER: US/09/820,924  
; CURRENT FILING DATE: 2001-03-30  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 39982  
; TYPE: DNA  
; ORGANISM: Human  
US-09-820-924-3



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2003, 21:01:54 ; Search time 7120.12 Seconds  
(without alignments)  
11491.278 Million cell updates/sec

Title: US-09-936-271b-13\_COPY\_1\_2000

Perfect score: 2000

Sequence: 1 gggccagagtgaggcaag.....gtgcgtctgcaccacacatc 2000

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl :

- 1: gb\_ba.\*
- 2: gb\_hg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*
- 15: em\_ba.\*
- 16: em\_fun.\*
- 17: em\_hum.\*
- 18: em\_in.\*
- 19: em\_mu.\*
- 20: em\_om.\*
- 21: em\_or.\*
- 22: em\_ov.\*
- 23: em\_pat.\*
- 24: em\_ph.\*
- 25: em\_pl.\*
- 26: em\_ro.\*
- 27: em\_sts.\*
- 28: em\_un.\*
- 29: em\_vi.\*
- 30: em\_hgt\_hum.\*
- 31: em\_hgt\_inv.\*
- 32: em\_hgt\_other.\*
- 33: em\_hgt\_mus.\*
- 34: em\_hgt\_pln.\*
- 35: em\_hgt\_rod.\*
- 36: em\_hgt\_mam.\*
- 37: em\_hgt\_vrt.\*
- 38: em\_sy.\*
- 39: em\_hgtgo\_hum.\*
- 40: em\_hgtgo\_mus.\*
- 41: em\_hgtgo\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	2000	100.0	11570	9	AF135028	AF135028 Homo sapi
C 2	1989	99.5	107487	9	AC011483	AC011483 Homo sapi
C 3	1989	99.5	230000	9	AF243527	AF243527 Homo sapi
C 4	1316.2	65.8	217346	2	AC027602	AC027602 Homo sapi
C 5	1227.2	61.4	200792	2	AC130782	AC130782 Pan trogl
C 6	1000	50.0	176647	2	AC130188	AC130188 Papio anu
C 7	461.6	23.1	32799	9	AC137591	AC137591 Homo sapi
C 8	456.6	22.8	141923	2	AC109997	AC109997 Homo sapi
C 9	455	22.8	138849	9	HS95C20	Z97181 Human DNA s
C 10	455	22.8	171697	9	AC114876	AC114876 Homo sapi
C 11	455	22.8	227137	9	AC098481	AC098481 Homo sapi
C 12	452.4	22.6	119147	9	AC098873	AC098873 Homo sapi
C 13	450.6	22.5	108893	9	AC079169	AC079169 Homo sapi
C 14	450	22.5	100208	9	AC135468	AC135468 Homo sapi
C 15	450	22.5	182569	9	AC092500	AC092500 Homo sapi
C 16	449.8	22.5	119555	9	BX004827	Human DNA
C 17	449.6	22.5	147122	9	AC090109	AC090109 Homo sapi
C 18	449	22.4	183690	9	AC091180	AC091180 Homo sapi
C 19	447.4	22.4	98876	9	HSJ90108	AL078461 Human DNA
C 20	444	22.2	246240	6	AR036572	AR036572 Sequence
C 21	444	22.2	246240	6	AR036573	AR036573 Sequence
C 22	444	22.2	246240	6	AR036574	U91328 Human hered
C 23	444	22.2	246282	9	HSU91328	AC068792 Homo sapi
C 24	443.4	22.2	106277	9	AC068792	AC068792 Homo sapi
C 25	442.4	22.1	101269	9	AC005300	AC005300 Homo sapi
C 26	442.4	22.1	193159	9	AC006946	AC006946 Homo sapi
C 27	442.4	22.1	196662	9	AP001107	AP001107 Homo sapi
C 28	442.4	22.1	198888	9	AC087728	AC087728 Papio anu
C 29	442	22.1	109907	9	AC092942	AC092942 Homo sapi
C 30	442	22.1	263876	9	AC019233	AC019233 Homo sapi
C 31	441.4	22.1	158003	9	AC022795	AC022795 Homo sapi
C 32	440.4	22.0	81278	9	AC091062	AC091062 Homo sapi
C 33	439.6	22.0	173058	9	AC104564	AC104564 Homo sapi
C 34	439.4	22.0	50054	9	AC093691	AC093691 Homo sapi
C 35	439	21.9	45896	9	HSBA9F11	AL050312 Human DNA
C 36	439	21.9	135640	9	AC036178	AC036178 Homo sapi
C 37	438	21.9	90007	9	AL139234	AL139234 Human DNA
C 38	438	21.9	95930	9	AC106033	AC106033 Homo sapi
C 39	438	21.9	166686	9	AC103895	AC103895 Papio anu
C 40	437.6	21.9	129711	9	AC110768	AC110768 Homo sapi
C 41	437.2	21.9	184396	9	AC010853	AC010853 Homo sapi
C 42	437.2	21.9	201861	9	AC067796	AC067796 Homo sapi
C 43	435.8	21.8	49816	9	AC093149	AC093149 Homo sapi
C 44	435.8	21.8	67199	9	AC025162	AC025162 Homo sapi
C 45	435.4	21.8	132720	9	AC093269	AC093269 Homo sapi

ALIGNMENTS

RESULT 1  
AF135028  
LOCUS AF135028 11570 bp DNA linear PRI 27-JUN-2000  
DEFINITION Homo sapiens kallikrein-like protein 2 KLK-L2 gene, complete cds.  
ACCESSION AF135028  
VERSION AF135028.1 GI:4589282  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 11570)  
AUTHORS Yousef, G.M. and Diamandis, E.P.  
TITLE The new kallikrein-like gene, KLK-L2. Molecular characterization,  
mapping, tissue expression, and hormonal regulation

J. Biol. Chem. 274 (53), 37511-37516 (1999)  
 20076408  
 MEDLINE  
 PUBMED  
 10608802  
 REFERENCE  
 2 (bases 1 to 11570)  
 Diamandis,E.P., Yousef,G.M., Luo,L.Y., Maqklara,A. and Obiezu,C.V.  
 The new human kallikrein gene family: implications in  
 carcinogenesis  
 Trends Endocrinol. Metab. 11 (2), 54-60 (2000)  
 21121728  
 MEDLINE  
 PUBMED  
 10675891  
 REFERENCE  
 3 (bases 1 to 11570)  
 Yousef,G.M., Luo,L.Y. and Diamandis,E.P.  
 Direct Submission  
 Submitted (13-MAR-1999) Pathology and Laboratory Medicine, Mount  
 Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,  
 Canada  
 FEATURES  
 source  
 Location/Qualifiers  
 1. 11570  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /map="19q13.3-q13.4"  
 join(2101..2131,2210..2293,4762..5023,5763..6019,  
 6105..6238,11092..11570)  
 /product="kallikrein-like protein 2 KLK-L2"  
 join(2221..2293,4762..5023,5763..6019,6105..6238,  
 11092..11247)  
 /codon\_start=1  
 /product="kallikrein-like protein 2 KLK-L2"  
 /protein\_id="AAD26429.1"  
 /db\_xref="GI:4589283"  
 /translation="MATARPWMVLCALITALLGVTEHVLANNVDSCHPSNTVPS  
 GSNQDLGAGEDARSDDSSRIINGSDCDHQTPWQAALLRPNQLYCAVLPHPW  
 LLTAHCKKRVFRLYGLSLSPVYESGGQFQVGRKIPHPHSGHSLNDLMLKLN  
 RRTPTKDVPIVNSSPCPSKGLVSGWGTTKSPQVFPKVLQCLINISVLQKRC  
 DAYPRQIDDTMFCAGDRGDSQSGGPPVVCNGLQGLVSWGIDYPCARNRPGVYT  
 NLCKFTKIQTQANS"  
 BASE COUNT 3034 a. 2562 c 3325 g 2649 t  
 ORIGIN  
 Query Match 100.0%; Score 2000; DB 9; Length 11570;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GGGCCAGAGTGAAGCAAGAGAGGAGTTGAGAGTCCCTCTCCTCAAGTGGCTTGATC 60  
 Db 1 GGGCCAGAGTGAAGCAAGAGAGGAGTTGAGAGTCCCTCTCCTCAAGTGGCTTGATC 60  
 Qy 61 TCCCTCGCTAAATGTCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
 Db 61 TCCCTCGCTAAATGTCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120  
 Qy 121 AAGAAAGAGAGAGAGAGAGAGACAGAAATACAACTACAGAAACACAGAGAGACAC 180  
 Db 121 AAGAAAGAGAGAGAGAGAGAGACAGAAATACAACTACAGAAACACAGAGAGACAC 180  
 Qy 181 ACAGAGAGCTGGGACACAGGACACAGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
 Db 181 ACAGAGAGCTGGGACACAGGACACAGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
 Qy 241 ACACAAATGGAGACACAGAGTGTAAAGAAAGAGAGAGATTAACAGAGTCCAGATACACGC 300  
 Db 241 ACACAAATGGAGACACAGAGTGTAAAGAAAGAGAGAGATTAACAGAGTCCAGATACACGC 300  
 Qy 301 AAAGGGCAGAGACAGAGTTTTCAGGTTGGTGTATGATCATCTCTTTTTTTTTTTT 360  
 Db 301 AAAGGGCAGAGACAGAGTTTTCAGGTTGGTGTATGATCATCTCTTTTTTTTTTTT 360  
 Qy 361 TTTTTTTTTTTTTTTGAGCGGAGCTCGCTCTGCTCGCCAGGCTGAGTCCAGTGGCGG 420  
 Db 361 TTTTTTTTTTTTTTTGAGCGGAGCTCGCTCTGCTCGCCAGGCTGAGTCCAGTGGCGG 420

Qy 421 GATCTCGGCTCACTGCAAGCTCCGCCCTCCGGGGTTACGGCATTTCTCTGCTCAGCCTC 480  
 Db 421 GATCTCGGCTCACTGCAAGCTCCGCCCTCCGGGGTTACGGCATTTCTCTGCTCAGCCTC 480  
 Qy 481 CCAAGTAGCTGGGACTACAGGCGCCGCCACTAGCCCGGGTAATTTTTTTTGTATTTT 540  
 Db 481 CCAAGTAGCTGGGACTACAGGCGCCGCCACTAGCCCGGGTAATTTTTTTTGTATTTT 540  
 Qy 541 GTAGAGAGGGGTTTACCGCTTTTAGCCGGGATGGCTCGATCTCCTGACCTCTGATCC 600  
 Db 541 GTAGAGAGGGGTTTACCGCTTTTAGCCGGGATGGCTCGATCTCCTGACCTCTGATCC 600  
 Qy 601 GCCCGCTCGCCCTCCCAAGTGTGAGGCTGATACAGGCTGAGCCACCGCCGCCCATG 660  
 Db 601 GCCCGCTCGCCCTCCCAAGTGTGAGGCTGATACAGGCTGAGCCACCGCCGCCCATG 660  
 Qy 661 ATCATCTTGTGACTATGCTGATGACAGTACTAAAGCATCAGACTCTACCCCTTTA 720  
 Db 661 ATCATCTTGTGACTATGCTGATGACAGTACTAAAGCATCAGACTCTACCCCTTTA 720  
 Qy 721 AATATGAGTTTGGCCAGGACCGTGGCTCATGCTGTAATTCAGCAGCTTTGGGAGGC 780  
 Db 721 AATATGAGTTTGGCCAGGACCGTGGCTCATGCTGTAATTCAGCAGCTTTGGGAGGC 780  
 Qy 781 AGAGTGGGTGAATCACTTGAAGGACGAGGTTTGAAGCCAGCTTGGCCACATGTTGAAA 840  
 Db 781 AGAGTGGGTGAATCACTTGAAGGACGAGGTTTGAAGCCAGCTTGGCCACATGTTGAAA 840  
 Qy 841 CTCTGCTTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 900  
 Db 841 CTCTGCTTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 900  
 Qy 901 ACCTGTAATCCAGCTATGCTGGAGGCTGAGGACGAGAGTCACTTGAACCTTGAGGCG 960  
 Db 901 ACCTGTAATCCAGCTATGCTGGAGGCTGAGGACGAGAGTCACTTGAACCTTGAGGCG 960  
 Qy 961 GAGGTTTCAGTGGGCGGAGATCATACACCGCCCTCCAGGCTGGGCGACAGAGCAAGCT 1020  
 Db 961 GAGGTTTCAGTGGGCGGAGATCATACACCGCCCTCCAGGCTGGGCGACAGAGCAAGCT 1020  
 Qy 1021 CTGCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1080  
 Db 1021 CTGCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1080  
 Qy 1081 AAAAAAATAATGCTGTCACAAATAGAGCAGAGTGAATAAAGGAAAAATAAATGGGCCA 1140  
 Db 1081 AAAAAAATAATGCTGTCACAAATAGAGCAGAGTGAATAAAGGAAAAATAAATGGGCCA 1140  
 Qy 1141 AGAACTCTAAGGTATATTGACAAATCATTGACAACTTTTAAAAAAGAAAGAACACAGA 1200  
 Db 1141 AGAACTCTAAGGTATATTGACAAATCATTGACAACTTTTAAAAAAGAAAGAACACAGA 1200  
 Qy 1201 GGCATAGAAAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
 Db 1201 GGCATAGAAAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
 Qy 1261 ACAAGGCTCTTAAGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
 Db 1261 ACAAGGCTCTTAAGACACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
 Qy 1321 AAAAAGACAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
 Db 1321 AAAAAGACAGAGAGAGAGACAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1380  
 Qy 1381 GAGAGAGGGTGGAGAGAGACAGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGGA 1440  
 Db 1381 GAGAGAGGGTGGAGAGAGACAGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGGA 1440  
 Qy 1441 GAACCCAGAGAGATGGAAGAAGACTCTGAGAAAAAACCCAGAGACAAAGATGGAAAGAGG 1500  
 Db 1441 GAACCCAGAGAGATGGAAGAAGACTCTGAGAAAAAACCCAGAGACAAAGATGGAAAGAGG 1500  
 Qy 1501 AGTATCGAGGGTGAACAGACAGTGGTGGAAATGAGCAAAATGCAGAGAGAAAGCAAGCAA 1560

Db 1501 AGTATCGAGGTGAACAGACAGTGTGGTAATGAGCAAAATGACAGAGAAAGCAAGCAA 1560  
QY 1561 TCCAGGCGCCAGCAATAGTAGCCAGAGTTGTGTGAGAACCCAGATCCTTAAGGCTGGGGG 1620  
Db 1561 TCCAGGCGCCAGCAATAGTAGCCAGAGTTGTGTGAGAACCCAGATCCTTAAGGCTGGGGG 1620  
QY 1621 AGGCAGGGAAGGGGTGGCTGGCTTCCGGAGACCCCTCCCATTTCTCCGGGCCAGGGAG 1680  
Db 1621 AGGCAGGGAAGGGGTGGCTGGCTTCCGGAGACCCCTCCCATTTCTCCGGGCCAGGGAG 1680  
QY 1681 GTAGGAGTGACATCCCGAGCTGGGTGGGGGTGTCTGGGGGTGGAGATAGGGGAGCA 1740  
Db 1681 GTAGGAGTGACATCCCGAGCTGGGTGGGGGTGTCTGGGGGTGGAGATAGGGGAGCA 1740  
QY 1741 GGAGGAGCTATTGCTTAAGGCCCGATAGGACACCTCATTTGCCCGGGATGTGCCCCAGGGAG 1800  
Db 1741 GGAGGAGCTATTGCTTAAGGCCCGATAGGACACCTCATTTGCCCGGGATGTGCCCCAGGGAG 1800  
QY 1801 CAGTGGGTGTTTAACTCAGGCCCGGTGCCAGAGCCCGAGGAGGAGGAGTGGCCAGGA 1860  
Db 1801 CAGTGGGTGTTTAACTCAGGCCCGGTGCCAGAGCCCGAGGAGGAGGAGTGGCCAGGA 1860  
QY 1861 AGGCACAGCCCTGAGAAAGTCTCGGCTGAGCTGGGAGCAAAATCCCCACCCCTACCTGG 1920  
Db 1861 AGGCACAGCCCTGAGAAAGTCTCGGCTGAGCTGGGAGCAAAATCCCCACCCCTACCTGG 1920  
QY 1921 GGCACAGGCCAAGTGAGACCTGGTGAGGTGGCTCAGCAGGAGGAGGAGGAGGTGTCT 1980  
Db 1921 GGCACAGGCCAAGTGAGACCTGGTGAGGTGGCTCAGCAGGAGGAGGAGGAGGTGTCT 1980  
QY 1981 GTGGCTCTCTGCACCCACATC 2000  
Db 1981 GTGGCTCTCTGCACCCACATC 2000

RESULT 2  
AC011483/c  
LOCUS 107487 bp DNA linear PRI 01-JUL-2002  
DEFINITION Homo sapiens chromosome 19 clone CTB-147C22, complete sequence.  
AC011483  
VERSION AC011483.7 GI:21637461  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 107487)  
Direct Submission  
Unpublished  
2 (bases 1 to 107487)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 107487)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
Direct Submission  
Submitted (01-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On Jul 1, 2002 this sequence version replaced gi:14971176.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.6% of Sequence;  
Estimated Total Number of Errors is 0.6.  
NOTE: Shatter libraries failed to resolve dinucleotide repeat.  
Unsure number of repeat copies 64998-65494. Forced join 65015.  
Location/Qualifiers  
1. .107487  
/organism="Homo sapiens"

/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTB-147C22"  
64998..65494  
/note="NOTE: Shatter libraries failed to resolve  
dinucleotide repeat. Unsure number of repeat copies  
64998-65494. Forced join 65015."  
BASE COUNT 26044 a 27304 c 25728 g 28411 t  
ORIGIN  
Query Match 99.5%; Score 1989; DB 9; Length 107487;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2000; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 1 GGGCCCCAGAGTGAAGCAGAGAGAGAGTTCAGAGCTCCCTCTGCAAAAGTGGCTTGAATC 60  
Db 60220 GGGCCCCAGAGTGAAGCAGAGAGAGAGTTCAGAGCTCCCTCTGCAAAAGTGGCTTGAATC 60161  
QY 61 TCCCTCGCTTAAATCAGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 119  
Db 60160 TCCCTCGCTTAAATCAGGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60101  
QY 120 GAAGAAAG 179  
Db 60100 GAAGAAAG 60041  
QY 180 CACAGAGAGCCCTGGGAGACAGAGGAGACACAGAGTTCAGAGAGAGAGAGAGAGAGAGAG 239  
Db 60040 CACAGAGAGCCCTGGGAGACAGAGGAGACACAGAGTTCAGAGAGAGAGAGAGAGAG 59981  
QY 240 GACACAAATGGAGACACAGAGGTGTAAAGAAAGAGAGAGATTAAACAGAGTCCACAGT 299  
Db 59980 GACACAAATGGAGACACAGAGGTGTAAAGAAAGAGAGATTAAACAGAGTCCACAGT 59921  
QY 300 CAAAGGGCAGAGACACAGTTCAGGGTGGTGTCTATGATCATCTCTCTTTTCTTTTCTTT 359  
Db 59920 CAAAGGGCAGAGACACAGTTCAGGGTGGTGTCTATGATCATCTCTCTTTTCTTTTCTTT 59861  
QY 360 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 419  
Db 59860 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 59801  
QY 420 GGATCTCGGCTCACATGCAAGCTCCGCTCCCGGTTTCAGCCATTCCTGCTCAGCCCT 479  
Db 59800 GGATCTCGGCTCACATGCAAGCTCCGCTCCCGGTTTCAGCCATTCCTGCTCAGCCCT 59741  
QY 480 CCCAAGTAGCTGGGACTACAGGCGCCCGCCACTACCGCGGCTAAATTTTCTTTTCTTTT 539  
Db 59740 CCCAAGTAGCTGGGACTACAGGCGCCCGCCACTACCGCGGCTAAATTTTCTTTTCTTTT 59681  
QY 540 ACTAGAGAGGGGTTCACCGTTTACCGGGATGCGCTCGATCTCCTGACCTCGTATC 599  
Db 59680 ACTAGAGAGGGGTTCACCGTTTACCGGGATGCGCTCGATCTCCTGACCTCGTATC 59621  
QY 600 CCGCCGCTCGGCTCCCAAGTGTCTGGGATACAGGCGTGGAGCCGACCGCCGCGCCAT 659  
Db 59620 CCGCCGCTCGGCTCCCAAGTGTCTGGGATACAGGCGTGGAGCCGACCGCCGCGCCAT 59561  
QY 660 GATCATCTTCTTTGACTATGCTGATGTGACAAGTACCTAAAGCCATCAGACTCTACCTTT 719  
Db 59560 GATCATCTTCTTTGACTATGCTGATGTGACAAGTACCTAAAGCCATCAGACTCTACCTTT 59501  
QY 720 AAATATGAGTTCGGCCAGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 779  
Db 59500 AAATATGAGTTCGGCCAGGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 59441  
QY 780 CAGAGTGGGTGAATCAGTTCAGGCGAGAGTTTGGAGACAGCTCGCCCAACATGCTGAA 839  
Db 59440 CAGAGTGGGTGAATCAGTTCAGGCGAGAGTTTGGAGACAGCTCGCCCAACATGCTGAA 59381  
QY 840 ACTCTGCTCTTACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 899  
Db 59380 ACTCTGCTCTTACTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 899

59380 ACTCTGCTTTTACTTAAAAA... 59321  
900 CACCTGTAATCCAGCTATGCTGAGGCTGAGGACGAGAGTCACTTGAACCTGGAGGC 959  
59320 CACCTGTAATCCAGCTATGCTGAGGCTGAGGACGAGAGTCACTTGAACCTGGAGGC 59261  
960 GGAGTTGTCAGTGGCGGAGATCATACCGCCCTCCAGCCTGGCGGAGAGCAAGAC 1019  
59260 GGAGTTGTCAGTGGCGGAGATCATACCGCCCTCCAGCCTGGCGGAGAGCAAGAC 59201  
1020 TCTGCTCAATAAATAAATAAACAAGCAAGCAGTGTGTTGTTACCTTATATATC 1079  
59200 TCTGCTCAATAAATAAATAAACAAGCAAGCAGTGTGTTGTTACCTTATATATC 59141  
1080 TAAAAAATAATGCTGTCAACAATAGCAGAGAGTGAATATAAGGAAATAAATGGGCC 1139  
59140 TAAAAAATAATGCTGTCAACAATAGCAGAGAGTGAATATAAGGAAATAAATGGGCC 59081  
1140 AAGAACTCTAAGGTATATTTGACAAATCATTCAGAACCTTTAAAAAAGAAGATCAG 1199  
59080 AAGAACTCTAAGGTATATTTGACAAATCATTCAGAACCTTTAAAAAAGAAGATCAG 59021  
1200 AGGCATAGAGACAGGAGGAGACAGGAGACAGAACACCTGTGGCCCAAGGAGACAA 1259  
59020 AGGCATAGAGACAGGAGGAGACAGGAGACAGAACACCTGTGGCCCAAGGAGACAA 58961  
1260 AACAGGCTCTTAAGACAGACAGGAGAGAGAGAGTGTGAGAGACAGACAGA 1319  
58960 AACAGGCTCTTAAGACAGACAGGAGAGAGAGAGTGTGAGAGACAGACAGA 58901  
1320 GAAAAAGACAGAGAGAGACAGACAGAGACAGAGAGCGGAGAGGGATAGAAAG 1379  
58900 GAAAAAGACAGAGAGAGACAGACAGAGAGAGAGAGCGGAGAGGGATAGAAAG 58841  
1380 AGCAGAGGGGTGAGAGACAGACAGATATTGAGAGAGACTCAGAAGATAGCCGAGG 1439  
58840 AGCAGAGGGGTGAGAGACAGACAGATATTGAGAGAGACTCAGAAGATAGCCGAGG 58781  
1440 AGAACACAGAGAGATGAAGAGACTCTGAGAAAAAACAGACAGACAAAGATGAAAGAG 1499  
58780 AGAACACAGAGATGAAGAGACTCTGAGAAAAAACAGACAGACAAAGATGAAAGAG 58721  
1500 GAGTATCGAGGTGAACAGACAGTGTGGAATAGCAAAATGACAGAGAAAGAAAGCA 1559  
58720 GAGTATCGAGGTGAACAGACAGTGTGGAATAGCAAAATGACAGAGAAAGAAAGCA 58661  
1560 ATCCAGGCGCAAGAATAGTACCCAGTGTGAGAGCCAGATCCTTAAGGCTGGG 1619  
58660 ATCCAGGCGCAAGAATAGTACCCAGTGTGAGAGCCAGATCCTTAAGGCTGGG 58601  
1620 GAGCAGGGAAGGGGCTGGCTTCCGAGAGACCCCTCCCAATTCCTCGGGCCAGGGA 1679  
58600 GAGCAGGGAAGGGGCTGGCTTCCGAGAGACCCCTCCCAATTCCTCGGGCCAGGGA 58541  
1680 GGTAGGAGTACATTCGGACTGGTGGGGGGTCTCTGGGGGTGAGATAGGGGGAGC 1739  
58540 GGTAGGAGTACATTCGGACTGGTGGGGGGTCTCTGGGGGTGAGATAGGGGGAGC 58481  
1740 AGGAGGAGCTATTGCTAAGGCGCCATAGGCACTTATGTCGGGGAATGTCGCCAGGGA 1799  
58480 AGGAGGAGCTATTGCTAAGGCGCCATAGGCACTTATGTCGGGGAATGTCGCCAGGGA 58421  
1800 GCAGTGGTGGTTTAACTCAGGCGCCGCTGCCAGAGCCCAAGGAGGAGTGGCCAGG 1859  
58420 GCAGTGGTGGTTTAACTCAGGCGCCGCTGCCAGAGCCCAAGGAGGAGTGGCCAGG 58361  
1860 AAGCAGAGGCTCAGAAAGTCTGCGCTGAGCTGGGAGCAATCCGCCACCCCTACCTG 1919  
58360 AAGCAGAGGCTCAGAAAGTCTGCGCTGAGCTGGGAGCAATCCGCCACCCCTACCTG 58301  
1920 GGGGACAGGCGCAAGTGTGAGGCTGTGAGGCTGTGCTCAGCAGGCGAGGAGAGGTTGTC 1979  
58300 GGGGACAGGCGCAAGTGTGAGGCTGTGAGGCTGTGCTCAGCAGGCGAGGAGAGGTTGTC 58241

QY 1380 TGTGCGTCTTGACCCACATC 2000  
Db 58240 TGTGCGTCTTGACCCACATC 58220

RESULT 3  
AF243527/c 230000 bp DNA linear PRI 21-NOV-2000  
LOCUS Homo sapiens serine protease gene cluster, complete sequence.  
DEFINITION AF243527  
ACCESSION AF243527  
VERSION AF243527.1 GI:11244757  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 230000)  
AUTHORS Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuall, J.,  
Moss, P., Paep, B. and Wang, K.  
TITLE Sequencing and expression analysis of the serine protease gene  
cluster located in chromosome 19q13 region  
JOURNAL Gene 257 (1), 119-130 (2000)  
MEDLINE 20510030  
PUBMED 11054574  
REFERENCE 2 (bases 1 to 230000)  
AUTHORS Gan, L., Lee, I., Smith, R., Argonza-Barrett, R., Lei, H., McCuall, J.,  
Moss, P., Paep, B. and Wang, K.  
TITLE Direct Submission  
JOURNAL Submitted (09-MAR-2000) Chiroscience R & D Inc., 1631 220th St. SE,  
Bothell, WA 98021, USA  
FEATURES  
source  
1. 230000  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13"  
complement(<963..>11517)  
/gene="KLK1"  
complement(join(<963..7118,7668..7804,7923..8212,  
9482..9641,11472..>11517))  
/gene="KLK1"  
/product="renal kallikrein"  
complement(join(6963..7118,7668..7804,7923..8212,  
9482..9641,11472..>11517))  
/gene="KLK1"  
/note="serine protease"  
/codon\_start=1  
/product="renal kallikrein"  
/protein\_id="AAG33353.1"  
/db\_xref="GI:11244758"  
/translation="MWFLVLCALSLGSGTAAPDQSRIVGGWECEHSPQWQAALVH  
FSFQCGILVHROWLTAHAICSDNYQLRLGRHNLFDDETAQFVHVSFPHPGN  
MSLLENTRQADEYSHDLMLRLTEPADTIDAVKVPPELPEPEVGSCTCAGSGS  
IPEFNFDLDQCVDLKILPNDCKRAHKQKVDYDFMLCVGHLEGGKDCVDSGGPL  
MCDGLQGVTSWGVPCGTPNPKPSAVRYSYVKWIEDTIAENS"  
complement(join(<13552..13704,14377..14530,14678..>14917,  
15416..15569,19204..>19246))  
/product="ACO protease"  
complement(join(13552..13704,14377..14530,14678..>14917,  
15416..15569,19204..>19246))  
/note="serine protease"  
/codon\_start=1  
/product="ACO protease"  
/protein\_id="AAG33354.1"  
/db\_xref="GI:11244759"  
/translation="MWLLTLFLSLASTAQDQKLLLEGDECAPHSQWQVALYERGR  
FNCGASLIPHWLAAHQCRRFMRVRLGHEHNLKRDGPOLRTTSTVIPPVYEARS  
HRNDIMLLRVLPARLPQVRPAVLPGPGEACVSGWGLVPLSPVSLPTLHC  
VNCIIISDTCKSYPGRLNTMVCAGAEGRGAECEGDSGGPLVCGGILQGIIVSWG  
VPDNTTKPGVTKVCHYLEWIRETKRN"  
complement(>42595..>47769)  
mRNA  
CDS  
gene

mrna	<p>/gene="Klk3" join(&lt;42395..42640,43880..44039,45669..45955,46099..46235,47614..&gt;47769) /gene="Klk3" /product="prostate specific antigen" join(42595..42640,43880..44039,45669..45955,46099..46235,47614..&gt;47769) /gene="Klk3" /note="serine protease" /codon_start=1 /product="prostate specific antigen" /protein_id="AAG33355.1" /db_xref="GI:11244760" /translation="MWVPVFLITLSVTWIGAAPLILSRIVGWCEKHSQPWQVLVASRGAVCGVLVHPQWLTAACHIRKNSVILLGRSLFHPEDTGVQVSHSFPPFLYDMSLKNRLSDSDSHDLMLRLSEPAELDAVKVMDLPTQEPALGTCYASGWSIPEEFLTPKKLQCDVLHVISNDVCAQVHPQVTFMLCAGRTGKSCSGDSGGPLVCNGVLQGITSGSEPCALPERPSLYTKVVRKWKIKDTIVANP" &lt;61139..&gt;66229 /gene="Klk2" join(&lt;61139..61184,62391..62550,64142..64428,64542..64678,66074..&gt;66229) /gene="Klk2" /product="glandular kallikrein" join(61139..61184,62391..62550,64142..64428,64542..64678,66074..&gt;66229) /gene="Klk2" /note="serine protease" /codon_start=1 /product="glandular kallikrein" /protein_id="AAG33356.1" /db_xref="GI:11244761" /translation="MWDLVLTSLSVGCTGAVPLTIQSRIVGWCEKHSQPWQVAVYSGHAGCGVLVHPQWLTAACHLKNQSVLGRHNLFPEDTGORVPVSHFPPFLYNMSLKHLSRPDESDSHDLMLRLSEPAKITDVVKVLGLTQEPALGTCYASGWSIPEEFLRPLQCSVLHLLSNDMCARAYSEKTFEMLCAGLTGKGTCCGDSGGPLVCNGVLQGITSGSEPCALPERPAVTVKVRKWKIKDTIVANP" complement(join(&lt;94602..94754,96027..96163,96247..96497,96919..97081,98345..&gt;98405)) /product="serine protease" complement(join(94602..94754,96027..96163,96247..96497,96919..97081,98345..98405)) /codon_start=1 /product="serine protease" /protein_id="AAG33357.1" /db_xref="GI:11244762" /translation="MATAGNPQWFLGVLILVAGSLVSGSCSIINGEDCSPHSQPQAALVWENELFCGVLVHPQWLSNAHCFQNSYITGLGHSLEADQEPGQWVNASLRVHPYPSLPLANDMLMLKLRNPTRPTDRPINVSSHCPSAGTKCLVSGWGTGKSPQVHFMTPLQVNVSVSEVCSKLYDPLHYPMFCAGGGHDKQDKSCNGDSGGPLVCNGVLQGITSGSEPCALPERPAVTVKVRKWKIKDTIVANP" complement(&lt;131301..&gt;140303) /gene="Klk5" /note="synonym: SCTE" complement(join(&lt;131301..131456,136310..136443,136529..136785,137525..137690,140255..&gt;140303)) /gene="Klk5" /product="stratum corneum trypsin-like serine protease" complement(join(131301..131456,136310..136443,136529..136785,137525..137690,140255..&gt;140303)) /gene="Klk5" /codon_start=1 /product="stratum corneum trypsin-like serine protease" /protein_id="AAG33358.1" /db_xref="GI:11244763" /translation="MWVLCALITALLGVTDARSDSSRIINGSDCDMHPTQWQAALLRPNQLYCGVLVHPQWLSNAHCFQNSYITGLGHSLEADQEPGQWVNASLRVHPYPSLPLANDMLMLKLRNPTRPTDRPINVSSHCPSAGTKCLVSGWGTGKSPQVHFMTPLQVNVSVSEVCSKLYDPLHYPMFCAGGGHDKQDKSCNGDSGGPLVCNGVLQGITSGSEPCALPERPAVTVKVRKWKIKDTIVANP" complement(join(&lt;146834..146986,149628..149764,151186..151433,155052..155208,155948..&gt;155987)) /product="protease M"</p>
cds	<p>/gene="Klk3" join(&lt;42395..42640,43880..44039,45669..45955,46099..46235,47614..&gt;47769) /gene="Klk3" /product="prostate specific antigen" join(42595..42640,43880..44039,45669..45955,46099..46235,47614..&gt;47769) /gene="Klk3" /note="serine protease" /codon_start=1 /product="prostate specific antigen" /protein_id="AAG33355.1" /db_xref="GI:11244760" /translation="MWVPVFLITLSVTWIGAAPLILSRIVGWCEKHSQPWQVLVASRGAVCGVLVHPQWLTAACHIRKNSVILLGRSLFHPEDTGVQVSHSFPPFLYDMSLKNRLSDSDSHDLMLRLSEPAELDAVKVMDLPTQEPALGTCYASGWSIPEEFLTPKKLQCDVLHVISNDVCAQVHPQVTFMLCAGRTGKSCSGDSGGPLVCNGVLQGITSGSEPCALPERPSLYTKVVRKWKIKDTIVANP" &lt;61139..&gt;66229 /gene="Klk2" join(&lt;61139..61184,62391..62550,64142..64428,64542..64678,66074..&gt;66229) /gene="Klk2" /product="glandular kallikrein" join(61139..61184,62391..62550,64142..64428,64542..64678,66074..&gt;66229) /gene="Klk2" /note="serine protease" /codon_start=1 /product="glandular kallikrein" /protein_id="AAG33356.1" /db_xref="GI:11244761" /translation="MWDLVLTSLSVGCTGAVPLTIQSRIVGWCEKHSQPWQVAVYSGHAGCGVLVHPQWLTAACHLKNQSVLGRHNLFPEDTGORVPVSHFPPFLYNMSLKHLSRPDESDSHDLMLRLSEPAKITDVVKVLGLTQEPALGTCYASGWSIPEEFLRPLQCSVLHLLSNDMCARAYSEKTFEMLCAGLTGKGTCCGDSGGPLVCNGVLQGITSGSEPCALPERPAVTVKVRKWKIKDTIVANP" complement(join(&lt;94602..94754,96027..96163,96247..96497,96919..97081,98345..&gt;98405)) /product="serine protease" complement(join(94602..94754,96027..96163,96247..96497,96919..97081,98345..98405)) /codon_start=1 /product="serine protease" /protein_id="AAG33357.1" /db_xref="GI:11244762" /translation="MATAGNPQWFLGVLILVAGSLVSGSCSIINGEDCSPHSQPQAALVWENELFCGVLVHPQWLSNAHCFQNSYITGLGHSLEADQEPGQWVNASLRVHPYPSLPLANDMLMLKLRNPTRPTDRPINVSSHCPSAGTKCLVSGWGTGKSPQVHFMTPLQVNVSVSEVCSKLYDPLHYPMFCAGGGHDKQDKSCNGDSGGPLVCNGVLQGITSGSEPCALPERPAVTVKVRKWKIKDTIVANP" complement(&lt;131301..&gt;140303) /gene="Klk5" /note="synonym: SCTE" complement(join(&lt;131301..131456,136310..136443,136529..136785,137525..137690,140255..&gt;140303)) /gene="Klk5" /product="stratum corneum trypsin-like serine protease" complement(join(131301..131456,136310..136443,136529..136785,137525..137690,140255..&gt;140303)) /gene="Klk5" /codon_start=1 /product="stratum corneum trypsin-like serine protease" /protein_id="AAG33358.1" /db_xref="GI:11244763" /translation="MWVLCALITALLGVTDARSDSSRIINGSDCDMHPTQWQAALLRPNQLYCGVLVHPQWLSNAHCFQNSYITGLGHSLEADQEPGQWVNASLRVHPYPSLPLANDMLMLKLRNPTRPTDRPINVSSHCPSAGTKCLVSGWGTGKSPQVHFMTPLQVNVSVSEVCSKLYDPLHYPMFCAGGGHDKQDKSCNGDSGGPLVCNGVLQGITSGSEPCALPERPAVTVKVRKWKIKDTIVANP" complement(join(&lt;146834..146986,149628..149764,151186..151433,155052..155208,155948..&gt;155987)) /product="protease M"</p>
mrna	<p>/gene="Klk3" join(&lt;42395..42640,43880..44039,45669..45955,46099..46235,47614..&gt;47769) /gene="Klk3" /product="prostate specific antigen" join(42595..42640,43880..44039,45669..45955,46099..46235,47614..&gt;47769) /gene="Klk3" /note="serine protease" /codon_start=1 /product="prostate specific antigen" /protein_id="AAG33355.1" /db_xref="GI:11244760" /translation="MWVPVFLITLSVTWIGAAPLILSRIVGWCEKHSQPWQVLVASRGAVCGVLVHPQWLTAACHIRKNSVILLGRSLFHPEDTGVQVSHSFPPFLYDMSLKNRLSDSDSHDLMLRLSEPAELDAVKVMDLPTQEPALGTCYASGWSIPEEFLTPKKLQCDVLHVISNDVCAQVHPQVTFMLCAGRTGKSCSGDSGGPLVCNGVLQGITSGSEPCALPERPSLYTKVVRKWKIKDTIVANP" &lt;61139..&gt;66229 /gene="Klk2" join(&lt;61139..61184,62391..62550,64142..64428,64542..64678,66074..&gt;66229) /gene="Klk2" /product="glandular kallikrein" join(61139..61184,62391..62550,64142..64428,64542..64678,66074..&gt;66229) /gene="Klk2" /note="serine protease" /codon_start=1 /product="glandular kallikrein" /protein_id="AAG33356.1" /db_xref="GI:11244761" /translation="MWDLVLTSLSVGCTGAVPLTIQSRIVGWCEKHSQPWQVAVYSGHAGCGVLVHPQWLTAACHLKNQSVLGRHNLFPEDTGORVPVSHFPPFLYNMSLKHLSRPDESDSHDLMLRLSEPAKITDVVKVLGLTQEPALGTCYASGWSIPEEFLRPLQCSVLHLLSNDMCARAY</p>



# JOURNAL REFERENCE AUTHORS

Unpublished  
2 (bases 1 to 217346)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
Boguslavsky,L., Boukhgalter,B., Brown,A., Burkett,G.,  
Campiolo,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Klein,J., LaRoque,K., Lamazares,R., Landers,T., Lehoczy,J.,  
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,  
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., MCPheeters,R.,  
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Menga,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (30-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 217346)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Beda,F., Boguslavsky,L.,  
Boukhgalter,B., Brown,A., Burkett,G., Campiolo,A., Castle,A.,  
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cooke,P.,  
Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Ferreira,P.,  
FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Goyette,M.,  
Graham,L., Grand-Pierre,N., Hagos,B., Heaford,A., Horton,L.,  
Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A., LaRoque,K.,  
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G.,  
Macdonald,P., Marquis,N., McCarthy,M., McEwan,P., McKernan,K.,  
MCPheeters,R., Meldrim,J., Meneus,L., Mihova,T., Menga,V.,  
Morrow,J., Murphy,T., Naylor,J., Norman,C.H., O'Connor,T.,  
O'Donnell,P., O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K.,  
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Rieback,M., Riley,R.,  
Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S., Severy,P.,  
Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
Tirrell,A., Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zimmer,A. and Zody,M.

Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Nov 15, 2000 this sequence version replaced gi:11136831.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L9166  
Center clone name: 795\_B\_6  
----- Summary Statistics  
Sequencing vector: M13; M7815; 31% of reads  
Sequencing vector: Plasmid; n/a; 69% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 210748 bases at least Q40  
Consensus quality: 213655 bases at least Q30  
Consensus quality: 215058 bases at least Q20  
Insert size: 194000; agarose-fp  
Insert size: 216246; sum-of-contigs  
Quality coverage: 11.9 in Q20 bases; agarose-fp

# JOURNAL REFERENCE AUTHORS

Quality coverage: 10.7 in Q20.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 8149: contig of 8149 bp in length  
\* 8150 8249: gap of 100 bp  
\* 8250 9592: contig of 1343 bp in length  
\* 9593 9692: gap of 100 bp  
\* 9693 10733: contig of 1041 bp in length  
\* 10734 10833: gap of 100 bp  
\* 10834 13519: contig of 2686 bp in length  
\* 13520 17510: gap of 100 bp  
\* 17511 17610: contig of 3891 bp in length  
\* 17611 24602: contig of 6992 bp in length  
\* 24603 24702: gap of 100 bp  
\* 24703 35434: contig of 10732 bp in length  
\* 35435 35534: gap of 100 bp  
\* 35535 124474: contig of 88940 bp in length  
\* 124475 124574: gap of 100 bp  
\* 124575 134564: contig of 10090 bp in length  
\* 134565 134764: gap of 100 bp  
\* 134765 162343: contig of 27579 bp in length  
\* 162344 162443: gap of 100 bp  
\* 162444 208917: contig of 46474 bp in length  
\* 208918 209017: gap of 100 bp  
\* 209018 217346: contig of 8329 bp in length.

FEATURES  
Location/Qualifiers  
1..217346  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19"  
/clone="RP11-795B6"  
/clone\_lib="RPC1-11 Human Male BAC"  
1..8149  
/note="assembly\_fragment"  
clone\_end:SP6  
vector\_side:left  
8250..9592  
/note="assembly\_fragment"  
9693..10733  
/note="assembly\_fragment"  
10834..13519  
/note="assembly\_fragment"  
13620..17510  
/note="assembly\_fragment"  
17611..24602  
/note="assembly\_fragment"  
24703..35434  
/note="assembly\_fragment"  
35535..124474  
/note="assembly\_fragment"  
124575..134564  
/note="assembly\_fragment"  
134765..162343  
/note="assembly\_fragment"  
162444..208917  
/note="assembly\_fragment"  
209018..217346  
/note="assembly\_fragment"  
clone\_end:17  
vector\_side:right  
BASE COUNT 55109 a 54154 c 53923 g 53053 t 1107 others  
ORIGIN

Query Match 65.8%; Score 1316.2; DB 2; Length 217346;

Best Local Similarity 95.4%; P-Val: 0.44e-314; Indels 5; Gaps 2;				
Matches 1343; Conservative 0; Mismatches 3;				
Qy	654	GGCCATGATCATCTCTTCTGACTATGCTGATGTGACAAAGTACCTTAAAGCCATCAGACTCTA	713	
Db	203789	GTCTATGATCATCTCTTCTGACTATGCTGATGTGACAAAGTACCTTAAAGCCATCAGACTCTA	203730	
Qy	714	CCCTTTAAATGATGAGTTTGGGCCAGGCACCGTGGCTCATGCTGTAATTCAGACATTT	773	
Db	203729	CCCTTTAAATGATGAGTTTGGGCCAGGCACCGTGGCTCATGCTGTAATTCAGACATTT	203670	
Qy	774	GGGAGGCAGAGGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGACAGCCTCGCCAAACAT	833	
Db	203669	GGGAGGCAGAGGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGACAGCCTCGCCAAACAT	203610	
Qy	834	GGTGAACCTGTCTTACTTAAATAAAAAAAAAAAAAAAAAAATCAGCCGGGTGCT	893	
Db	203609	GGTGAACCTGTCTTACTTAAATAAAAAAAAAAAAAAAAAAATCAGCCGGGTGCT	203551	
Qy	894	GGGCGACACCTGTATATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTGAACCT	953	
Db	203550	GGGCGACACCTGTATATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTGAACCT	203491	
Qy	954	GGAGCGGAGGTCGACTGGCGCGAGATCATCACCGCCTCCAGCCTGGGCGACAGAG	1013	
Db	203490	GGAGCGGAGGTCGACTGGCGCGAGATCATCACCGCCTCCAGCCTGGGCGACAGAG	203431	
Qy	1014	CAAGACTCTGCTCAAAATAAATAAACAACAGAACAGCAGTGTGTACCTTAGT	1073	
Db	203430	CAAGACTCTGCTCAAAATAAATAAACAACAGAACAGCAGTGTGTACCTTAGT	203371	
Qy	1074	TATATCTAAAAAATAATGCTGTCAACAATAGACGAGAAGTCGAATTAAGGAAATAAA	1133	
Db	203370	TATATCTAAAAAATAATGCTGTCAACAATAGACGAGAAGTCGAATTAAGGAAATAAA	203311	
Qy	1134	TGGGCCAAGAAGCTCTTAGGTATATTTCACAAATCATTCAGAACCTTTAAAAAAGAAAGAA	1193	
Db	203310	TGGGCCAAGAAGCTCTTAGGTATATTTCACAAATCATTCAGAACCTTTAAAAAAGAAAGAA	203251	
Qy	1194	TCACAGAGGCATAGAAAGACAGAGGAGGAGAACAGGAGAGACAGAAACACTGTGGCCCAAGGA	1253	
Db	203250	TCACAGAGGCATAGAAAGACAGAGGAGGAGAACAGGAGAGACAGAAACACTGTGGCCCAAGGA	203191	
Qy	1254	GAAACAAACAGGCTCTTAACACAGACAGGAG --- -	1309	
Db	203190	GAAACAAACAGGCTCTTAACACAGACAGGAGGAGAGAGAGAGAGAGTGTGAGTGAGA	203131	
Qy	1310	GACACACAGAGAAAGACACAGAGAGAGACACAGACAGACAGAGAGAGAGCGCAGAG	1369	
Db	203130	GACACACAGAGAAAGACACAGAGAGAGAGACACAGACAGACAGAGAGAGCGCAGAG	203071	
Qy	1370	GGATAGAAGAGAGAGAGGGGTGGAGAGACACGAGATATTGAGAGAGACTCAGAAAGA	1429	
Db	203070	GGATAGAAGAGAGAGAGGGGTGGAGAGACACGAGATATTGAGAGAGACTCAGAAAGA	203011	
Qy	1430	TAGCCGAGGGAGAACCCAGAGAGATGGAAGAGCTCTGAGAAAAAACCCAGAGACAAAG	1489	
Db	203010	TAGCCGAGGGAGAACCCAGAGAGATGGAAGAGCTCTGAGAAAAAACCCAGAGACAAAG	202951	
Qy	1490	ATGGAAGAGAGAGTATCCAGGGGTGAACACAGACAGTGGTGGAAATGAGCAAAAATGCAGAGAAG	1549	
Db	202950	ATGGAAGAGAGTATCCAGGGGTGAACACAGACAGTGGTGGAAATGAGCAAAAATGCAGAGAAG	202891	
Qy	1550	AAAGCAAGCAATCCAGGCGCCAGAAATAGTAGTACCAGTGTGGTGAAGCCAGATCCCTT	1609	
Db	202890	AAAGCAAGCAATCCAGGCGCCAGAAATAGTAGTACCAGTGTGGTGAAGCCAGATCCCTT	202831	
Qy	1610	AAAGCTGGGGAGGAGGAGGAGGCTGGCTGGCTTCCGAGAGACCCCTCCCAATTCCTC	1669	
Db	202830	AAAGCTGGGGAGGAGGAGGAGGCTGGCTGGCTTCCGAGAGACCCCTCCCAATTCCTC	202771	
Qy	1670	GGGCCAGGAGGATAGGAGTAGTACATTCGGAGCTGGGTGGGGGGTGTCTCTGGGGGTGGAGA	1729	

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring

clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8x average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a phrap-derived quality score.

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 19656 base at least Q40  
Consensus quality: 197983 bases at least Q30  
Consensus quality: 198679 bases at least Q20  
Insert size: 215000; agarose- $\phi$   
Insert size: 199692; sum-of-contigs  
Quality coverage: 9.73x in Q20 bases; agarose- $\phi$   
Quality coverage: 9.73x in Q20 bases; sum-of-con-

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* been provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

*	1	28306:	contig	of 28306 bp	in length
*	28307	28406:	gap	of unknown length	
*	28407	37836:	contig	of 9450 bp	in length
*	37857	37952:	gap	of unknown length	
*	37957	73522:	contig	of 35566 bp	in length
*	73523	73622:	gap	of unknown length	
*	73623	83567:	contig	of 9945 bp	in length
*	83568	83667:	gap	of unknown length	
*	83668	88817:	contig	of 5150 bp	in length
*	88818	88917:	gap	of unknown length	
*	88918	125611:	contig	of 36694 bp	in length
*	125612	125711:	gap	of unknown length	
*	125712	159879:	contig	of 34168 bp	in length
*	159880	159979:	gap	of unknown length	
*	159980	174698:	contig	of 14719 bp	in length
*	174699	174798:	gap	of unknown length	
*	174799	186382:	contig	of 11584 bp	in length
*	186383	186482:	gap	of unknown length	
*	186483	193344:	contig	of 6862 bp	in length
*	193345	193444:	gap	of unknown length	
*	193445	199363:	contig	of 5919 bp	in length
*	199364	199463:	gap	of unknown length	
*	199464	200792:	contig	of 1329 bp	in length

## FEATURES

```

misc feature
1..28306
location/qualifiers
1..200792
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="CH251-355A20"
/clone_lib="CH251"

```

misc feature

misc\_feature

misc\_feature

**misc\_feature**

misc\_feature

misc\_feature

misc\_feature

[illegible]

Query Match	61.4%	Score 1227.2;	DB 2;	Length 200792;
Best Local Similarity	95.5%	Pred. No. 3.4e-292;		
Matches 1323;	Conservative	0;	Mismatches 18;	Indels 45; Gaps 4;
Qy	654	GGCAGTATCATCTTCTTGACTATGCTGATGTGACAAGTACCTAAAGCCATCAGACTCTA	713	
Db	78497	GTCTATGATCATCTTCTTGACTATGCTGATGTGACAAGTACCTAAAGCCATCAGACTCTA	78438	
Qy	714	CCCTTTAAATPATGCAGTTTGGCCAGGCCCGTGCCTCATGCTGTAATTCAGACATTT	773	
Db	78437	CCCTTTAAATPATGCAGTTTGGCCAGGCCACTGTGGCTCATGTATTAATTCAGCACATTT	78378	
Qy	774	GGGAGCGAGGTGGGTGAATCATCTTAGCCACGAGTTTGAGCCAGCCTGGCCCAACAT	833	
Db	78377	GGGAGCGAGGTGGGTGAATCATCTTAGCCACGAGTTTGAGCCAGCCTGGCCCAACAT	78318	
Qy	834	GGTGAACCTCTGCTTTACTTAAAAAATAAAAAAAAAAAAAAAAAATCAGCCGGTGTGCT	893	
Db	78317	GGTGAACCTCTGCTTTACTT-----AAAAAATAAAAAAAAAATCAGCCGGTGTGCG	78264	
Qy	894	GGGCGACACCTGTAAATCCACGTAATGCTGGAGGCTGAGGCACGAGATCACTTGAACCT	953	
Db	78263	GGGCGACACCTGTAAATCCACGTAATGCTGGAGGCTGAGGCACGAGATCACTTGAACCT	78204	
Qy	954	GGAGCGGAGGTTGTCAGTGGCGCGAGATCACATCACCGCCCTCCAGCCTGGCGACACAG	1013	
Db	78203	GGAGGTGGAGTTGTCAGTGGCGCGAGATCACATCACCGCCCTCCAGCCTGGCGACACAG	78144	
Qy	1014	CAAGACTCTGCTCAAAATAAATAAATAAACAAGCAAGCAGTTTGTGTACCTTAGT	1072	
Db	78143	CAAGACTCTGCTCAAAATAAATAAATAAACAAGCAAGCAGTTTGTGTACCTTAGT	78084	
Qy	1074	TATATCT--AAAAAATAATGCTGTCAACAAATAGACGAGAAGTGAATAAAGGAAATA	1131	
Db	78083	TATACCTTAAAAAATAATGCTGTCAACAAATAGACGAGAAGTGAATAAAGGAAATA	78024	
Qy	1132	AATGGGCCAAGAACTCTAAGTGATATTTCACAATATCTTCAGAACCTTTAAAAAAGAAG	1191	
Db	78023	AATGGGCCAAGAACTCTAAGTGATATTTCACAATATCTTCAGAACCTTTAAAAAAGAAG	77964	
Qy	1192	AATCAGAGGGCATAGAAAGACAGGGAGAAACAGGGAGACAAACCTGTGGCCCAAG	1251	
Db	77963	AATCAGAGGGCATAGAAAGACAGGGAGAAACAGGGAGACAAACCTGTGGCCCAAG	77904	
Qy	1252	GAGAAACAAACAGGCTCTTAAGACAGACAGGAG-----GAGAGAGAGAGAGTCACTGA	1307	
Db	77903	GAGAAACAAACAGGCTCTTAAGACAGACAGGAGGGCGAGAGAGAGAGAGTCACTGA	77844	
Qy	1308	GAGACACAGAGAAAGAACACAGAGAGAGAGACAGAGA-----1347		
Db	77843	GAGACACAGAGAAAGAACACAGAGAGAGAGACAGAGAGCGGAGGGGATAGAAAG	77784	
Qy	1348	-----CAGAGACAGAGAGCGGAGGGATAGAAAGACAGAGAGGGGTAGAAAG	1394	
Db	77783	AGAGAGGGGTTGGAGAGACAGAGAGCGGAGGGATAGAAAGAGAGAGAGGGGTGA	77724	
Qy	1395	GAGAGACACAGATATTGAGAGAGACTCAGAAGATAGCCGAGGGAGAACACACAGAGA	1454	

||||| 77723 GAGAGACGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGAGAAACACACAGAGA 77664  
QY 1455 TGGAGAGACTCTGAGAAAAACCCAGACAAAGATGGAAGAGGAGTATCGAGGGTGA 1514  
Db 77663 TGGAGAGAGACTCTGAAAAAAACCCAGAGCAAGATGGAAGAGGAGTATCGAGGGTGA 77604  
QY 1515 ACACACAGTGTGGAATGAGCAAAATGAGAGAAAGCAAGCAATCCAGGCGCCCAAGA 1574  
Db 77603 ACACACAGTGTGGAATGAGCAAAATGAGAGAAAGCAAGCAATCCAGGCGCCCAAGA 77544  
QY 1575 ATAGTACCCAGAGTGTGTGAGAGCCAGATCTTAAGGCTTGGGGGAGGAGCGAAGGGG 1634  
Db 77543 ATAGTGACCCAGAGTGTGTGAGAGCCAGATCTTAAGGCTTGGGGGAGGAGCGAAGGGG 77484  
QY 1635 CTGGCTCGCTTCGGAGACCCCTCCCATTTCTCCGGCCAGGAGGTAGGAGTGACAT 1694  
Db 77483 CTGGCTCGCTTCCAGAGACCCCTCCCATTTCTCCGGCCAGGAGGTAGGAGTGACAT 77424  
QY 1695 TCCGACTGGGTGGGGGTGCTCTGGGGTGGAGATAGGGGAGCAGGAGGAGCTATTGC 1754  
Db 77423 TCCGACTGGGTGGGGGTGCTCTGGGGTGGAGATAGGGGAGCAGGAGGAGCTATTGC 77364  
QY 1755 TAAGCCCGATAGGACCTATTGCCCCGGGAATGTGCCAGGAGAGCAGTGGGTGGTTAT 1814  
Db 77363 TAAGCCCGATAGGACCTATTGCCCCGGGAATGTGCCAGGAGAGCAGTGGGTGGTTAT 77304  
QY 1815 AACTCAGGCCCGTCCAGAGCCAGGAGGAGCAGTGCCAGGAGGACAGGACAGCCCTGA 1874  
Db 77303 AACTCAGGCCCGTCCAGAGCCAGGAGGAGCAGTGCCAGGAGGACAGGACAGCCCTGA 77244  
QY 1875 GAAGTCTCGGCTGAGCTGGGAGCAAAATCCCCACCCCTACCTGGGGGAGCAGGCAAGT 1934  
Db 77243 GAAGTCTCGGCTGAGCTGGGAGCAAAATCCCCACCCCTACCTGGGGGAGCAGGCAAGT 77184  
QY 1935 GAGACCTGGTGGGTGCTCAGCAGGAGGAGGAGAGGTCTGTGCGCTCTGTCACC 1994  
Db 77183 GAGACCTGGTGGGTGCTCAGCAGGAGGAGGAGAGGTCTGTGCGCTCTGTCACC 77124  
QY 1995 CACATC 2000  
|||||  
Db 77123 CACATC 77118

RESULT 6  
AC130188/c  
LOCUS AC130188 176647 bp DNA linear HTG 14-NOV-2002  
DEFINITION Papio anubis clone RP41-421P3, WORKING DRAFT SEQUENCE, 12 ordered  
pieces.  
ACCESSION AC130188  
VERSION AC130188.2 GI:24960890  
KEYWORDS HTG; HTGS\_PHASE2; HTGS\_DRAFT.  
SOURCE Papio anubis (olive baboon)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
Cercopitheciinae; Papio.  
1 (bases 1 to 176647)  
Ahter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S. M.,  
Benjamin, B., Blakesley, R. W., Bouffard, G. G., Brinkley, C., Brooks, S.,  
Carla, K., Coleman, B., Engle, J., Granlitz, S., Guan, X., Gupta, J.,  
Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Idol, J. R., Karlins, E.,  
Laric, P., Lee-Lin, S.-Q., Legaspi, R., Maduro, O. L., Maduro, V. B.,  
Marquies, E. H., Mastello, C., Maskeri, B., McDowell, J.,  
Paquirigan, C., Pearson, R., Portnoy, M. E., Prasad, A.,  
Redix-Dugue, N., Schandler, K., Schueler, M. G., Sison, C.,  
Stantripo, S., Thomas, J. W., Thomas, P. J., Touchman, J. W., Vogt, J. L.,  
Wetherby, K. D., Wiggins, L., Young, A. and Green, E. D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 176647)  
Green, E. D.  
Direct Submission

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

Submitted (08-AUG-2002) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA  
3 (bases 1 to 176647)  
Green, E. D.  
Direct Submission  
Submitted (14-NOV-2002) NIH Intramural Sequencing Center, 8717  
Grovemont Circle, Gaithersburg, MD 20877, USA  
On Nov 14, 2002 this sequence version replaced gi:22138439.  
----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: http://www.nisc.nih.gov  
Contact: nisc\_zoo@nhgri.nih.gov  
----- Project Information  
Center project name: dsr  
Center clone name: 421P03

The sequence data in this record represents an 'enhanced'  
version of a Phase 2 submission. Specifically, the indicated  
order and orientation of each sequence contig has been  
established using one or more of the following: read-pair  
data from individual subclones, overlaps with neighboring  
clones, alignment with available reference sequence (e.g.,  
human), and/or confirmation by PCR testing. In addition,  
the sequence assembly is based on at least 8x average  
coverage in Q20 bases and has been reviewed to rule out  
gross misassemblies, the low-quality ends of sequence  
contigs have been trimmed away, and each base is associated  
with a Phrap-derived quality score.

----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 171127 bases at least Q40  
Consensus quality: 173424 bases at least Q30  
Consensus quality: 174803 bases at least Q20  
Insert size: 188000; agarose-fp  
Insert size: 175547; sum-of-contigs  
Quality coverage: 8.76x in Q20 bases; agarose-fp  
Quality coverage: 9.39x in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 12 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.

\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

\* 1  
6526: contig of 6526 bp in length  
6527: gap of unknown length  
6627: contig of 12186 bp in length  
18813: gap of unknown length  
18913: contig of 42868 bp in length  
61781: gap of unknown length  
97068: contig of 35188 bp in length  
97069: gap of unknown length  
97169: contig of 28038 bp in length  
125207: gap of unknown length  
125207: contig of 1715 bp in length  
127022: gap of unknown length  
127121: contig of 2266 bp in length  
129387: gap of unknown length  
129487: contig of 2815 bp in length  
129488: gap of unknown length  
132303: contig of 3333 bp in length  
132403: gap of unknown length  
135735: contig of 7723 bp in length  
135835: gap of unknown length  
143558: contig of 26524 bp in length  
143559: gap of unknown length  
170182: contig of 6365 bp in length  
170183: gap of unknown length  
170283: contig of 6365 bp in length



SOURCE  
ORGANISM

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

1 (bases 1 to 32799)  
Rump, A., Rosenthal, A., Drescher, B. and Schattevoy, R.  
Direct Submission  
Submitted (07-MAY-1997) Genome Analysis, Institute Of Molecular  
Biotechnology, Beutenbergstrasse 11, Jena 07745, Germany

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

2 (bases 1 to 32799)  
Lagemann, D. and Platzter, M.  
Direct Submission  
Submitted (26-NOV-2002) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

3 (bases 1 to 32799)  
Lagemann, D. and Platzter, M.  
Direct Submission  
Submitted (05-DEC-2002) Genome Analysis, Institute of Molecular  
Biotechnology, Beutenbergstr. 11, Jena, Thuringia 07745, Germany

## COMMENT

-----  
Center: Institute of Molecular Biotechnology  
Center code: IMB  
Web site: <http://genome.imb-jena.de/>  
Contact: gscj-submit@genome.imb-jena.de  
-----  
Project Information  
-----  
Center project name: B6  
Center clone name: XX-540F24  
-----  
Summary Statistics  
Sequencing vector: pUC18; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 32765 bases at least Q40  
Consensus quality: 32770 bases at least Q30  
Consensus quality: 32770 bases at least Q20  
Quality coverage: 13.12x  
-----

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one pUC18 subclone.

## Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

## FEATURES

## source

-----  
Location/Qualifiers  
1..32799  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/map="xp22-PAR"  
/clone="XX-B6cos"  
537..642  
/note="single stranded/single chemistry region"  
2807..2820  
/note="single stranded/single chemistry region"  
2913..2927  
/note="low quality region"  
3166..3173  
/note="single stranded/single chemistry region"  
5131..5234  
/note="single stranded/single chemistry region"  
5235..5317  
/note="single clone coverage"  
-----

## misc\_feature

5266..5317  
/note="low quality region"

## unsure

5266..5317  
5366..5431  
/note="single clone coverage"

## misc\_feature

5432..5518  
/note="single stranded/single chemistry region"

## misc\_feature

6151..6395  
/note="single stranded/single chemistry region"

## unsure

6303  
6382  
unsure

## misc\_feature

7106..7137  
/note="single stranded/single chemistry region"

## misc\_feature

9265..9528  
/note="single stranded/single chemistry region"

## variation

9669  
/note="T substituted in clone: XX-p99cos"

## variation

10317  
/replace="C"

## variation

10897  
/note="deleted in clone: XX-540F24"

## variation

11087..11090  
/replace="T"

## variation

11087..11090  
/note="deleted in clone: XX-540F24"

## variation

11259  
/replace="T"

## misc\_feature

11744..11756  
/note="low quality region"

## misc\_feature

11766..11880  
/note="single stranded/single chemistry region"

## variation

11834  
/note="T substituted in clone: XX-540F24"

## variation

11892..11900  
/note="deleted in clone: XX-1BCos"

## variation

12197..12200  
/note="deleted in clone: XX-1BCos"

## variation

12256..12259  
/note="deleted in clone: XX-1BCos"

## variation

12277  
/note="A substituted in clone: XX-540F24"

## variation

12290..12291  
/note="deleted in clone: XX-540F24"

## variation

12312  
/note="G substituted in clone: XX-540F24"

## variation

13362  
/note="G substituted in clone: XX-1BCos"

## variation

13768  
/note="T substituted in clone: XX-540F24"

## variation

14138  
/note="T substituted in clone: XX-540F24"

## variation

14664  
/note="C substituted in clone: XX-540F24"

## variation

15355  
/note="deleted in clone: XX-1BCos"

## variation

15447..15448  
/note="deleted in clone: XX-1BCos"

variation	15481. .15483 /note="deleted in clone: XX-540F24" /replace=""	15888	QY	625	TGGGATTACAGCGCTGAGCCACCGCGCCGCGCATGATCATCTTCTTACTATCTGCTGATG 684 
variation	/note="A substituted in clone: XX-540F24" /replace="G"	15961	Db	6112	TGGGATTACAGCGCTGAGCCACCGCGCGCATGATCATCTTCTTAAACATTTGTAG 6053 
variation	/note="C substituted in clone: XX-540F24" /replace="T"	16252	QY	685	TGACAAAGTACCTAAAGCCATCAGACTCTACCCCTTTAAATATGAGTTTGGCCGAGGCACC 744 
variation	/note="C substituted in clone: XX-540F24" /replace="T"	16334	Db	6052	ATTAAATTAAAGTGTGCGCGTGTTTAGA-----AACTGGCCTCTTGGCCAGGCACA 5999 
variation	/note="A substituted in clone: XX-1Bcos" /replace="G"	16684	QY	745	GTGGCTCATCCCTGTAATTCACGACTTTGGGAGGAGAGGTGGGTGAATCACTTGTAGGC 804 
variation	/note="C substituted in clone: XX-540F24" /replace="A"	16793	Db	5998	GTGGCTCATCCCTGTAATTCACGACTTTGGGAAATTTGAGGTGGGCGAAATCAC--GAGGT 5941 
variation	/note="G substituted in clone: XX-540F24" /replace="A"	17129	QY	805	CAGGAGTTTGAGACCAGCCTGGCCACATGCTGAACTCTGTCTTTTACTATAAAAAA 864 
variation	/note="C substituted in clone: XX-540F24" /replace="T"	17489	Db	5940	CAGGAATTCAGACCGCTGGCCACACACCGTGAACCCCATCTCTACTA----- 5891 
variation	/note="C substituted in clone: XX-540F24" /replace="G"	17812. .18717	QY	865	AAAAAATAAATAAATCAGCGGTGTGTGGGCGACACCTGTATATCCAGCTATGTCTGGA 924 
variation	/note="G substituted in clone: XX-540F24" /replace="A"	21509. .21686	Db	5890	--AATATACAAAAATAGCTGGGAGTGGTGGCGCATGCTTGTATATCCAGCACTTGGGA 5833 
variation	/note="C substituted in clone: XX-540F24" /replace="G"	21628	QY	925	GGCTGAGGCAGAGAGTCACCTTGAACCCCTGGAGCGGAGGTGCAGTGGCCGAGATCAC 984 
variation	/note="C substituted in clone: XX-540F24" /replace="T"	22529	Db	5832	GGCTGAGGCAGAGAGTCGCTTGAACCCAGGAGGTGGAGGTGCATTTGAGCCGAGATGCG 5773 
variation	/note="G substituted in clone: XX-540F24" /replace="A"	22559. .22586	QY	985	ATCACCGCCTCCAGCCTGGCGGACAGAGCAAGACTCTGTCTCAATATAATAAACA 1044 
variation	/note="C substituted in clone: XX-540F24" /replace="G"	22606	Db	5772	GCCTCTGCACTCCAGCCT--GGTACAGAGCAAGACTCT-----TAAAGAGAGGAGG 5720 
misc_feature	/note="single stranded/single chemistry region"	22606	QY	1045	AACGAACAAGCAGTGTGTGTACCTTATATCTTAAATAAATAAATAAATAAATAAATAA 1104 
misc_feature	/note="single stranded/single chemistry region"	22606	Db	5719	AAGGAAGGAAGAGAGAGGAG 5660 
variation	/note="A substituted in clone: XX-540F24" /replace="G"	22606	QY	1105	AGAGCAGAGTGAATAAAGGAATAAATAAATGGGCCAGAGACTCTAAGGTATATTTGACAA 1164 
variation	/note="G substituted in clone: XX-540F24" /replace="A"	22606	Db	5659	AAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 5601 
variation	/note="G substituted in clone: XX-540F24" /replace="A"	22606	QY	1165	ATCATTCAGAAACCTTTAAAAAAGAAAGAAATCACAGAGGCTATAGAAAGACAGAGGAG 1224 
variation	/note="G substituted in clone: XX-540F24" /replace="A"	22606	Db	5600	GAAAG 5541 
misc_feature	/note="single stranded/single chemistry region"	22606	QY	1225	GGGAGACAGAAACCTGTGTG--CCCAGGAGAGACAAACAAAGGCTCTCTAGACAGACAGG 1283 
misc_feature	/note="single clone coverage"	22606	Db	5540	AGGAGAAAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 5481 
misc_feature	/note="low quality region"	22606	QY	1284	AGGAGAGAGAGAGAGAGTGTGAGTGTGAGACAGACAGAGAGAGAGAGAGAGAGAGAGAG 1343 
misc_feature	/note="low quality region"	22606	Db	5480	GGAGGGAAGGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 5421 
misc_feature	/note="G substituted in clone: XX-540F24" /replace="A"	22606	QY	1344	GAGACAGAGAGAGAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1402 
misc_feature	/note="G substituted in clone: XX-540F24" /replace="A"	22606	Db	5420	TGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 5361 
misc_feature	/note="G substituted in clone: XX-540F24" /replace="A"	22606	QY	1403	CGAGATATTGAGAGAGACTCAGAAAGATAGCCGAGGAGAGAGAGAGAGAGAGAGAGAGAG 1462 
misc_feature	/note="G substituted in clone: XX-540F24" /replace="A"	22606	Db	5360	AAGAAGAAATGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 5301 
misc_feature	/note="G substituted in clone: XX-540F24" /replace="A"	22606	QY	1463	GACTCTGAGAAAAACACAGAGACAAAGATGGAAGAGAGTATGTGAGGTGTGACAGACAG 1522 
misc_feature	/note="G substituted in clone: XX-540F24" /replace="A"	22606	Db	5300	GAAAGAAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 5241 
misc_feature	/note="G substituted in clone: XX-540F24" /replace="A"	22606	QY	1523	TGTTGGAATTCAGCAAAATTCAG 1582 
misc_feature	/note="G substituted in clone: XX-540F24" /replace="A"	22606	Db	5240	AGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 5181 
misc_feature	/note="G substituted in clone: XX-540F24" /replace="A"	22606	QY	1583	CCAGAGTTGTGTGAGAGCCAGATCCTTTAAGGCTGGGGGAGGCGAGGAGGAGGAGG 1634 
misc_feature	/note="G substituted in clone: XX-540F24" /replace="A"	22606	Db	5180	GGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGAAGGA 5129 

RESULT 8  
AC109997  
LOCUS  
DEFINITION Homo sapiens chromosome 5 clone CTD-207601, WORKING DRAFT SEQUENCE,

AC109997 141923 bp DNA linear HTG 09-FEB-2002





## Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program. All manually edited bases have been reduced to quality zero. Quality levels above 40 are expected to have less than 1 error in 10,000 bp. Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., Phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

## Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below. The electronically digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC. Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

## BgIII

## HindIII

## EcoRI

SeqDerMap	FngPrnt	SeqDerMap	FngPrnt	SeqDerMap	FngPrnt
2292	2308	6857	6790	8696	8687
2067	2005	6382	6450	6	<800
4763	4915	512	<800	3403	3464
2585	2578	449	<800	334	<800
1981	2005	7940	7924	951	930
5496	5398	10618	10372	4919	4873
8296	8303	305	<800	3694	3709
10488	10186	1052	1054	5455	5487
5655	5680	2795	2867	555	<800
6535	6574	1527	1521	13221	13066
2695	2731	9418	9389	8640	8687
10234	10186	2970	3015	5654	5681
4871	4915	10593	10372	2842	2844
1822	1781	891	883	10800	10593
726	720	1004	1054	11874	11745
540	<800	1340	1314	8529	8687
11711	11427	817	883	286	<800
25420	25600	4767	4687	6660	6724
2222	2308	7571	7564	13999	13754

360	<800	4492	4508	9092	9112
2344	2308	1092	1054	2356	2380
1198	1180	3332	3396	2287	2380
3828	3831	1123	1054	2669	2713
3328	3300	1803	1824	924	930
3248	3300	3041	3122	4641	4650
1289	1252	16514	16912	1247	1206
399	<800	192	<800	1708	1662
4134	4173	4763	4687	6082	6133
900	895	587	<800	1904	1897
3852	3963	5086	5041	4433	4394
4463	4378	2278	2421	6851	6941
3959	3831	1336	1314	1799	1783
1990	2005	3223	3122	23885	24086
4272	4173	7004	7057	-----	-----
2327	2308	971	1054	-----	-----
6858	6866	2341	2421	-----	-----
3056	3063	1649	1646	-----	-----
6815	6866	2822	2867	-----	-----
809	810	6084	6048	-----	-----
4433	4378	163	<800	-----	-----
6135	6032	9648	9389	-----	-----
-----	-----	311	<800	-----	-----
-----	-----	3279	3396	-----	-----
-----	-----	3900	3950	-----	-----
-----	-----	2035	2029	-----	-----
-----	-----	3612	3630	-----	-----
-----	-----	3053	3229	-----	-----
-----	-----	3414	3396	-----	-----
-----	-----	865	883	-----	-----
-----	-----	637	<800	-----	-----
-----	-----	259	<800	-----	-----
-----	-----	97	<800	-----	-----
-----	-----	1242	1314	-----	-----
-----	-----	270	<800	-----	-----
-----	-----	23	<800	-----	-----
-----	-----	47	<800	-----	-----





```

QY 746 TGCTCATGCTGTAATTCAGCACATTGGAGGCGAGAGTGGTGAATCACTTGAGGCC 805
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 91862 TGCTCAGCGCTGTAATCCAGCACATTGGAGGCGAGAGTGGTGAATCACTTGAGGCC 91921
QY 806 AGAGATTGAGACCGAGCTGGCCCAACAGTGTGAACCTGCTCTTTACTTAAAAA 865
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 91922 CGAGTTCGAGACAGCTGGCCCAACAGTGTGAACCTGCTCTTTACTTAAAAA 91969
QY 866 AAAAAAATAATCAGCGGCTGCTGGGCGACACCTGTAATCCAGCTATGCTGGAG 925
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 91970 AAAAAAATAATCAGCGGCTGCTGGGCGACACCTGTAATCCAGCTATGCTGGAG 92029
QY 926 GCTGAGGCGAGGAGTCACTTGAACCTGAGCGGAGGTGTCAGTGGCGGAGATCACA 985
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92030 TCTGAGGCGAGGAGTCACTTGAACCTGAGCGGAGGTGTCAGTGGCGGAGATCACA 92089
QY 986 TCACCGCCCTCAGCTGGCGGAGAGCAAGTCTGCTCAATAAATAAATAACAA 1045
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 92090 CCACTGTACTCCAGCCTGGTGGAGCAAGACTCTGTCTCAAAAAATAAATAAATAA 92149
QY 1046 ACGAA 1050
    || ||
Db 92150 AATAA 92154

```

## RESULT 12

```

AC098873
LOCUS Homo sapiens BAC clone RP11-793H20 from 4, complete sequence.
AC098873 AC036224
VERSION AC098873.3 GI:18855169
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 119147)
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
PUBMED 9847074
REFERENCE 2 (bases 1 to 119147)
AUTHORS Walligorski, J., Haakenson, W. and Spalding, L.
TITLE The sequence of Homo sapiens BAC clone RP11-793H20
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 119147)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 119147)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (21-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 119147)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (09-MAR-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Feb. 21, 2002 this sequence version replaced gi:18042516.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0793H20
Drafting Center: WIBR

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

## SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E., Taten, M., Catanese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBACE3.6

## NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-260K18, 2000 bp overlap; the clone sequenced to the right is RP11-473N2. Actual start of this clone is at base position 106515 of RP11-260K18; actual end is at base position 119147 of RP11-793H20.

Single plasmid region exists between 77265 and 77307. Polymorphisms exist between AC079118, AC024191 and AC098873.

The sequence of AC036224 has been incorporated into AC098873.

FEATURES	source	Location/Qualifiers
	1. .119147	/organism="Homo sapiens"
		/mol_type="genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="4"
		/map="4"
repeat_region		/clone="RP11-793H20"
	167..273	/clone.lib="RPCI-11"
repeat_region		/rpt_family="MaLR"
	841..887	/rpt_family="MaLR"
misc_feature		/rpt_family="T-rich"
	870..889	/note="match to EST BG114275 (NID:gl2607872)"
repeat_region		/rpt_family="L2"
	2009..2357	/rpt_family="L2"
repeat_region		/rpt_family="MaLR"
	2369..2934	/rpt_family="MaLR"
repeat_region		/rpt_family="L2"
	2959..3655	/rpt_family="L2"
repeat_region		/rpt_family="MIR"
	4049..4289	/rpt_family="MIR"
repeat_region		/rpt_family="MIR"
	4448..4566	/rpt_family="MIR"
repeat_region		/rpt_family="MIR"
	4908..4933	/rpt_family="MIR"
repeat_region		/rpt_family="(TGAA)n"
	4936..5061	/rpt_family="(TGAA)n"
repeat_region		/rpt_family="MER103"
	5574..5879	/rpt_family="MER103"
repeat_region		/rpt_family="Alu"
	5949..6101	/rpt_family="Alu"



## VERSION

AC079169.32 GI:21263154

## KEYWORDS

HTG. Homo sapiens (human)

## SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## ORGANISM

Mammalia; Euthera; Primates; Catarrhini; Hominoidea; Homo.

## REFERENCE

## AUTHORS

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Worley, K.C.

Direct Submission

Submitted (22-JAN-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On May 30, 2002 this sequence version replaced gi:21206031.

INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email

gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

## ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished) for Human and Mouse sequences.

Genes and region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT:

## FEATURES

## Location/Qualifiers

1..108893  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/clone="RP11-47911"

## misc\_feature

1..2067  
/note="overlaps bases 108750..110816 of clone AC002404"

## repeat\_region

/function="clone overlap"  
complement(549..580)

## repeat\_region

/rpt\_family="LIM4"  
complement(581..871)

## repeat\_region

/rpt\_family="AluX"  
complement(872..1369)

## repeat\_region

/rpt\_family="LIM4"  
1370..1394

## repeat\_region

/rpt\_family="(TTTTG)n"  
complement(1398..1675)

## repeat\_region

/rpt\_family="AluX"  
complement(1760..2236)

## repeat\_region

/rpt\_family="LIM4b"  
2259..2373

## repeat\_region

/rpt\_family="AluJb"  
complement(2375..2450)

## repeat\_region

/rpt\_family="LIM4b"  
complement(2472..2666)

## repeat\_region

/rpt\_family="LIMC/D"  
2683..2963

## TITLE

## JOURNAL

## AUTHORS

## COMMENT

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

Db	50324	GAATGGTCTCGAATCTCCTGACCTGCTGATCCGCCGCCCTCGCGCTCCCAAGTGTCTGGGA	50383
QY	630	TTACAGCGGTGAGCCACCGCGCCCGCCATGATCATCTTCTTGACTATGCTGATGTGACA	689
Db	50384	TTACAGCGGTGAGCCACCGCGCCCGCCCTGCTTAAGTCTTTTATATGTTCTTGTGGAG	50443
QY	690	AGTACATAAGCCATCAGACTCTACCCCTTTAAATATGCAGTTTGGCCAGCAGCGTGGC	749
Db	50444	AATATAT-TATTCAACATTCACCTTGTAATAGGAAATATATTTAGCCAGGAACAGTGGC	50502
QY	750	TCATGCTGTAATTCAGCAGCTTTGGGAGGCGAGAGTGGTGAATCACTTGGAGCCAGGA	809
Db	50503	TCGCGCTTGTAATCCAGACACTTTGGGGGCGAGAGCGAGATCACTGAGTFCAGGT	50562
QY	810	GTTTGTAGACGAGCTGGCCAACTGTGTAACTCTGTCTTTACTAAAAAATAAATAAATAA	869
Db	50563	GTTTGTAGACGAGCTGGCCAACTGTGTAACTCTGTCTTTACTAAAAAATAAATAAATAA	50610
QY	870	AAAAAATAATCAGCGGGGTGCTGTGGGGGCACACCTGTAAATCCAGCTATGCTGGAGGCTG	929
Db	50611	ATACAAAAATTAGCTGGCGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	50670
QY	930	AGGCAGAGAGTCACTTTGAACCTGTGGAGGCGAGGTTGCAGTGGGCCAGAGATCACATCAC	989
Db	50671	AGGCAGAAGAATCACCTTTGAACCCAGGAGGAGGTTGCAGTGGGCCAGAGATCACATCAC	50730
QY	990	CGCCCTCCAGCCTGGCGGCGAGCAGCAAGACTCTGCTCAAAATAAATAAATAAATAA	1044
Db	50731	TGCACCTCCAGCCTGGGTGACCGAGATTCGGTCTCAAAAATAAATAAATAAATAAATAA	50785
RESULT 14			
AC135468			
LOCUS	AC135468	100208 bp	DNA linear PRI 17-MAY-2003
DEFINITION	Homo sapiens chromosome 1 clone RP11-298P9, complete sequence.		
ACCESSION	AC135468 AL450462		
VERSION	AC135468.3 GI:30841055		
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and Haugen, E.D.		
TITLE	Direct Submission		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 100208)		
AUTHORS	Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (16-OCT-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA		
REFERENCE	3 (bases 1 to 100208)		
AUTHORS	Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and Haugen, E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-NOV-2002) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA		
REFERENCE	4 (bases 1 to 100208)		
AUTHORS	Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z., Saenphimmachak, C., Buckley, D., Kibukawa, M., Raymond, C. and Haugen, E.D.		
TITLE	Direct Submission		
JOURNAL	Submitted (17-MAY-2003) Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA		
COMMENT	On May 17, 2003 this sequence version replaced gi:25705321.		
	----- Genome Center		
	Center: University of Washington Genome Center		
	Center Code: UWGC		
	Web site: <a href="http://www.genome.washington.edu">http://www.genome.washington.edu</a>		
	Contact: <a href="mailto:uwgchq@u.washington.edu">uwgchq@u.washington.edu</a>		

## Drafting Center: SC

----- Project Information  
 Center project name: chr-1  
 Center clone name: sc0392  
 sc0392 was derived from the original BAC (RP11-298P9)  
 using yeast homologous capture technique.

## ----- Summary Statistics

Sequencing vector: plasmid; 53% of reads  
 Sequencing vector: plasmid; L08752; 47% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 100206 bases at least Q40  
 Consensus quality: 100207 bases at least Q30  
 Consensus quality: 100208 bases at least Q20  
 Insert size: 100208; sum-of-contigs  
 Quality coverage: 25.1x in Q20 bases; sum-of-contigs

## Overlapping Sequences:

5': RP11-297123 (UWGC:sc0384) AC098656, 3777-bp overlap  
 3': RP5-871E2 AL161797, 3773-bp overlap

## Sequence Quality Assessment:

This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.  
 All manually edited bases have been reduced to quality zero.  
 Quality levels above 40 are expected to have less than 1 error in 10,000 bp.  
 Base-by-base quality values are not generally visible from the Genbank flat file format but are available as part of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:  
 all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest.

## Sequence Validation:

This sequence has been validated by Multiple Complete Digest fingerprinting. Comparison of the experimentally derived digest fragments with sequence-predicted fragments is given below.  
 The electronically-digested sequence consists of both insert and vector, in order to accurately represent the entire circular BAC.  
 Small fragments below a variable cutoff (approximately 400-800 bp) are not resolved in the fingerprint and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragments are separated by dashed lines.

HindIII				EcoRI				BglII			
SeqDerMap	FngrPnt	SeqDerMap	FngrPnt	SeqDerMap	FngrPnt	SeqDerMap	FngrPnt	SeqDerMap	FngrPnt	SeqDerMap	FngrPnt
27861	27870	2775	2822	4460	4442						
10835	10615	15690	15691	2067	2066						
8	<800	5046	4994	4849	4849						
1191	1145	593	<800	2389	2395						
541	<800	121	<800	8936	8933						
10071	10047	4176	4218	3336	3347						
7336	7338	5633	5550	859	854						
818	843	4670	4557	24	<800						
3168	3194	4598	4557	29	<800						

8217	8355	1084	1065	1752	1748
7535	7745	77	<800	1057	1047
477	<800	766	776	297	<800
5023	5020	619	<800	10235	10177
61	<800	1381	1379	1580	1550
4008	4017	459	<800	336	<800
1890	1924	1679	1655	2820	2839
5242	5141	2384	2388	3407	3423
846	843	1873	1835	5948	5883
3480	3460	1659	1655	835	854
823	843	4721	4758	603	<800
2118	2082	3389	3384	4452	4442
1996	1924	1331	1379	6616	6652
2474	2515	2045	2050	1944	1944
2746	2745	1252	1237	10220	10177
		3321	3384	7959	7933
		1441	1379	3562	3548
		3094	3079	3091	3117
		3442	3471	9174	9315
		169	<800	5473	5486
		1055	1065	445	<800
		3927	3866		
		1288	1313		
		4579	4557		
		5480	5550		
		41	<800		
		5950	6031		
		2655	2688		
		15	<800		
		213	<800		
		432	<800		
		688	<800		
		172	<800		
		73	<800		
		2709	2745		

FEATURES  
 source  
 1..100208 Location/Qualifiers







GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2003, 20:39:54 ; Search time 493.736 Seconds  
(without alignments)  
10934.739 Million cell updates/sec

Title: US-09-936-271B-13\_COPY\_1\_2000

Perfect score: 2000

Sequence: 1 gggccagagtgaggcaag.....gtgctctgcaccacatc 2000

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_19Jun03.\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.\*
- 25: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2000	100.0	11570	21	Human KLK-L2 gene.
C 2	424.4	21.2	84607	20	Human PACAP genom
C 3	416.6	20.8	139904	24	Human CDNA differe
4	415.2	20.8	17397	22	Human cardiovascul
5	415.2	20.8	19334	22	Human cardiovascul
6	415.2	20.8	19345	22	Human cardiovascul
C 7	414.8	20.7	12970	22	Human reproductive
C 8	414.8	20.7	12970	23	Human testicular a

C 9	414.6	20.7	74037	24	ABK94412	DNA encoding endot
10	410.4	20.5	10445	22	AA103692	Human reproductive
11	406.2	20.3	12970	22	AA105001	Human reproductive
12	406.2	20.3	12970	23	ABL97894	Human testicular a
C 13	406	20.3	32224	22	AAK89986	Human digestive sy
14	405.8	20.3	4433	22	ABAI8400	Human nervous syst
C 15	405.8	20.3	4433	22	AA540516	DNA encoding human
C 16	405.8	20.3	4433	22	AA104142	Human reproductive
C 17	404.6	20.2	33147	22	AAK67282	Human immune/haema
C 18	403.6	20.2	1503841	24	ABT00010	Human neuregulin 1
C 19	403.6	20.2	1503841	24	ABT01503	Human neuregulin 1
C 20	403.6	20.2	1503900	22	AAK95240	Human neuregulin-1
C 21	403.6	20.2	1503900	22	AAK96733	Human neuregulin-1
C 22	402.6	20.1	36785	22	AAK82208	Human immune/haema
C 23	401.8	20.1	4431	22	ABA18399	Human nervous syst
C 24	401.8	20.1	4431	22	AA540513	DNA encoding human
C 25	401.8	20.1	4431	22	AA104139	Human reproductive
C 26	401.6	20.1	23241	22	AAF97870	Human neuroblastom
C 27	401.6	20.1	23241	22	AAF97871	Human neuroblastom
C 28	400.6	20.0	17717	22	AAK82007	Human immune/haema
29	400.6	20.0	17717	22	AAK82008	Human immune/haema
30	400.6	20.0	17719	22	AAK82006	Human immune/haema
31	400	20.0	19038	22	AAF30492	Human PAK5 gene.
C 32	398.4	19.9	25971	22	AAK86336	Human immune/haema
C 33	396.6	19.8	37783	22	AAK70780	Human immune/haema
C 34	396.6	19.8	37783	22	AAK76625	Human immune/haema
C 35	396.6	19.8	37783	22	AAK80913	Human immune/haema
C 36	394.8	19.7	4033	22	AAF90327	Human JAFFA genom
C 37	394.4	19.7	3608	22	AAK67271	Human immune/haema
C 38	394.4	19.7	3608	22	AAK74891	Human immune/haema
C 39	394.4	19.7	3608	22	AAK83192	Human immune/haema
40	394.4	19.7	20247	22	AA136315	Human musculoskele
41	394.4	19.7	20247	25	ABX59303	CDNA encoding nove
C 42	393.4	19.7	109906	24	ABK94411	DNA encoding endot
C 43	393	19.7	139904	24	ABK83562	Human cDNA differe
C 44	392.8	19.6	222930	24	ABK84349	Human cDNA differe
C 45	391.4	19.6	19199	22	AAK70995	Human immune/haema

ALIGNMENTS

RESULT 1  
AAA95905  
ID AAA95905 standard; DNA; 11570 BP.  
XX  
AC AAA95905;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE Human KLK-L2 gene.  
XX  
XX Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;  
KW kallikrein-like protein; serine protease; cytosolic; cancer;  
KW prostrate cancer; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200053776-A2.  
XX  
PD 14-SEP-2000.  
XX  
PF 09-MAR-2000; 2000WO-CA00258.  
XX  
PR 11-MAR-1999; 99US-0124260.  
PR 01-APR-1999; 99US-0127386.  
PR 21-JUL-1999; 99US-0144919.  
PA (MOUN ) MOUNT SINAI HOSPITAL.  
XX  
PI Yousef GM, Diamandis EP;  
XX  
DR WPI; 2000-587440/55.

DR P-PSDB: AAB21296.  
XX New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L  
PT protein mediated disorders, especially cancer.  
XX Claim 1; Page 143-149; 184pp; English.  
XX The present sequence is the coding sequence of the human KLK-L2 gene,  
CC which encodes a kallikrein-like protein. Kallikreins and kallikrein-like  
CC proteins are a subgroup of the serine protease enzyme family. They  
CC catalyze the selective cleavage of specific polypeptide precursors to  
CC release peptides with potent biological activity. Nucleic acids encoding  
CC kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and  
CC KLK-L6 have been isolated. The proteins are useful in the treatment,  
CC monitoring and diagnosis of cancers, especially prostate cancer. They  
CC can also be used to identify a substance that can associate with or  
CC mediate the biological activity of the proteins. Antibodies can be used  
CC to treat conditions mediated by the kallikrein-like proteins.  
XX SQ Sequence 11570 BP; 3033 A; 2562 C; 3326 G; 2649 T; 0 other;  
Query Match 100.0%; Score 2000; DB 21; Length 11570;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2000; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGCCAGAGTCAAGGCAAGAGAGAGTTCAGAGCTCCCTCTGCAAGTGGCTTGAGTC 60  
DB 1 GGGCCAGAGTCAAGGCAAGAGAGAGTTCAGAGCTCCCTCTGCAAGTGGCTTGAGTC 60  
QY 61 TCCCTGCTAAATGTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
DB 61 TCCCTGCTAAATGTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120  
QY 121 AAGAAAG 180  
DB 121 AAGAAAG 180  
QY 181 ACAG 240  
DB 181 ACAG 240  
QY 241 ACACAAATGGAG 300  
DB 241 ACACAAATGGAG 300  
QY 301 AAAGGGCAG 360  
DB 301 AAAGGGCAG 360  
QY 361 TTTTCTTTTCTTTTGTAG 420  
DB 361 TTTTCTTTTCTTTTGTAG 420  
QY 421 GATCTGCGCTCACTGCAAGTCCCGCTCCCGGTTTCCAGGCTTCTCTGCTGCTGCTGCTG 480  
DB 421 GATCTGCGCTCACTGCAAGTCCCGCTCCCGGTTTCCAGGCTTCTCTGCTGCTGCTGCTG 480  
QY 481 CCAAGTACTGGAGTCAAG 540  
DB 481 CCAAGTACTGGAGTCAAG 540  
QY 541 GTAG 600  
DB 541 GTAG 600  
QY 601 GCCCGCTGCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 660  
DB 601 GCCCGCTGCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCGCTCCCG 660  
QY 661 ATCATCTTCTGACTATCTGATGTGACAGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAG 720  
DB 661 ATCATCTTCTGACTATCTGATGTGACAGTAAAGGAGAGAGAGAGAGAGAGAGAGAGAG 720

QY 721 AATATGAGTGTGGCCAGGACCGTGGCTCATCCCTGTATATCCAGCACTTTGGGAGGC 780  
DB 721 AATATGAGTGTGGCCAGGACCGTGGCTCATCCCTGTATATCCAGCACTTTGGGAGGC 780  
QY 781 AGAGTGGTGAATCACTTGAGGCCAGGAGTTTTCAGACCAGCTTGGCCAAATGTTGAA 840  
DB 781 AGAGTGGTGAATCACTTGAGGCCAGGAGTTTTCAGACCAGCTTGGCCAAATGTTGAA 840  
QY 841 CTCTGCTTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 900  
DB 841 CTCTGCTTTTACTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 900  
QY 901 ACCTGTATATCCAGCTATGCTGGAGGCTAGGCGAGAGAGTCACTTGAACCTTGAGGCG 960  
DB 901 ACCTGTATATCCAGCTATGCTGGAGGCTAGGCGAGAGAGTCACTTGAACCTTGAGGCG 960  
QY 961 GAGCTTCAGTGGCGGAGATCAGATCAGCCCTCCAGCTGGGCGAGAGAGAGAGAGAG 1020  
DB 961 GAGCTTCAGTGGCGGAGATCAGATCAGCCCTCCAGCTGGGCGAGAGAGAGAGAGAG 1020  
QY 1021 CTGCTCAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1080  
DB 1021 CTGCTCAATATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1080  
QY 1081 AAAAAAATAATGCTGCAACAAATAGAGCAGAGTGAATAAAGAAATAAATGAGGCGCA 1140  
DB 1081 AAAAAAATAATGCTGCAACAAATAGAGCAGAGTGAATAAAGAAATAAATGAGGCGCA 1140  
QY 1141 AGAAGTCTAGGTATATTTGACAAATCATTCAGAACCTTTTAAAAAAGAAAGAAATCA 1200  
DB 1141 AGAAGTCTAGGTATATTTGACAAATCATTCAGAACCTTTTAAAAAAGAAAGAAATCA 1200  
QY 1201 GGCATAGAAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
DB 1201 GGCATAGAAAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260  
QY 1261 ACAAGGCTCTTAAGACAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
DB 1261 ACAAGGCTCTTAAGACAGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320  
QY 1321 AAAAGACAG 1380  
DB 1321 AAAAGACAG 1380  
QY 1381 GAGAGAGGCTGAGAGAGAGAGAGAGAGATATTTAGAGAGAGTCAAGAAAGATACCGG 1440  
DB 1381 GAGAGAGGCTGAGAGAGAGAGAGAGAGATATTTAGAGAGAGTCAAGAAAGATACCGG 1440  
QY 1441 GAACACAGAGAGATGGAAGAGAGTCTGTAGAAAAAACCCAGAGACAAAGATGGAAGAG 1500  
DB 1441 GAACACAGAGAGATGGAAGAGAGTCTGTAGAAAAAACCCAGAGACAAAGATGGAAGAG 1500  
QY 1501 AGTATCAGGCTGAACAGAGAGAGTGGTGAATGAGCAAAATTCAGAGAGAGAGAGAGAG 1560  
DB 1501 AGTATCAGGCTGAACAGAGAGAGTGGTGAATGAGCAAAATTCAGAGAGAGAGAGAGAG 1560  
QY 1561 TCCAGGCGCCAAAGATAGTACCCAGAGTGGTGTAGAGAGCCAGATCTCTTAAGGCTGGGG 1620  
DB 1561 TCCAGGCGCCAAAGATAGTACCCAGAGTGGTGTAGAGAGCCAGATCTCTTAAGGCTGGGG 1620  
QY 1621 AGGAGGAGAGGAGGCTGGCTGGGTTCCGGAGAGACCCCTCCCATCTTCCGGGCGGAGAG 1680  
DB 1621 AGGAGGAGAGGAGGCTGGCTGGGTTCCGGAGAGACCCCTCCCATCTTCCGGGCGGAGAG 1680  
QY 1681 GTAGGAGTGAATTCGGAGTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGG 1740  
DB 1681 GTAGGAGTGAATTCGGAGTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGGTTGGG 1740  
QY 1741 GGAGGAGTATTTGAGGCGCCGATAGGAGTCACTTATTCGCCGGAATGTGCCAGGAGAG 1800  
DB 1741 GGAGGAGTATTTGAGGCGCCGATAGGAGTCACTTATTCGCCGGAATGTGCCAGGAGAG 1800  
QY 1801 CAGTGGTGGTATTAATCAAGGCGGCTGGCCAGAGAGGAGGAGGAGGAGGAGGAGGAG 1860





KW Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;  
KW antihemetic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebrotrophic; nootropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; tissue regeneration;  
KW anti-infertility.  
XX Homo sapiens.  
XX WO200155321-A2.  
XX PD 02-AUG-2001.  
XX PF 17-JAN-2001; 2001WO-US01340.  
XX 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249246.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249266.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.



PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226729.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-02321968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 01-NOV-2000; 2000US-02441826.  
PR 08-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.

PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0251988.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-451930/48.

XX New cardiovascular system related polynucleotides and polypeptides,  
PT useful for diagnosing, treating and/or preventing disorders of the  
PT cardiovascular system -

XX Claim 1; SEQ ID NO 1943; 674pp; English.

XX Sequences AAS35741-AAS36942 represent genomic DNA molecules, which encode  
CC the cardiovascular system antigen polypeptides of the invention.

CC Cardiovascular system antigens and their associated polynucleotides are  
CC useful in the diagnosis, treatment and prevention of various types of  
CC disorders in e.g. humans, mice, rabbits, goats, cats, dogs,  
CC chickens or sheep. A pathological condition can be determined by  
CC detecting the presence or absence of a mutation in a cardiovascular  
CC system antigen polynucleotide. The treatable disorders include autoimmune  
CC diseases such as rheumatoid arthritis, hyperproliferative disorders such  
CC as neoplasms of the breast or liver, cardiovascular disorders such as  
CC cardiac arrest, cerebrovascular disorders such as cerebral ischaemia,  
CC nervous system disorders such as Alzheimer's disease, infections caused  
CC by bacteria, viruses and fungi, ocular disorders such as corneal

CC infection, endocrine disorders such as premature labour and infertility,  
CC gastrointestinal disorders such as Crohn's disease, renal disorders such  
CC as glomerulonephritis and respiratory disorders such as asthma and  
CC pleurisy. The polypeptides can also be used to aid wound healing, to  
CC prevent skin aging due to sunburn, to maintain organs before  
CC transplantation, to regenerate tissues and in chemotaxis.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX

Query Match 20.8%; Score 415.2; DB 22; Length 19334;  
Best Local Similarity 78.6%; Pred. No. 5.1e-64;  
Matches 558; Conservative 0; Mismatches 123; Indels 29; Gaps 4;

```
Qy 349 TTTTNTTTTTTTTTTTTTTTTTTTTTTTTGTGACAGGAGTCTGCTGCGCCAGGTGGA 408
Db 15143 TGTTCGTAACCTCTTTTTGTGTTTAAAGACAGACTTGTGCTGCGCCAGGTGGA 15202

Qy 409 GTGCAGTGGGGGAGTCTCGGCTCACTGCAAGCTCCCGCTCCCGGGTTTCAGGCCATTCTCC 468
Db 15203 GTGCAGTGGGACGATCTCGGCTCACTGCAAGCTCCCGTCTCCGGGTTTCAGGCCATTCTCC 15262

Qy 469 TGCCTCAGCTCCCAAGTAGTGGGACTACAGGGCCCGGCACTACGCCCGGGCTAATTTT 528
Db 15263 TGCCTCAGCTCCCGAGTAGTGGGACTACAGGCACCCGCGCACCAAGCCGGGTAA-TTT 15321

Qy 529 TTTGTATTTTGTAGTAGAGGGGTTTACCGGTTTGTAGCCGGGATGCCCTGATCTCCTG 588
Db 15322 TTTGTATTTTGTAGTAGAGGGGTTTGTCCGTTGTAGCCAGGATGCTCGATCTCCTG 15381

Qy 589 ACCTCGTGATCCCGCCCGCTCCGCTCCCAAGTGTGGGATTACAGCGGTGAGCCACCG 648
Db 15382 ACCTGTGATCTGCGCCCTCGGCTCCCAAGTGTGGGATTACAGCGGTGAGCTACTG 15441

Qy 649 CGCCCGGCGCATGATCTTCTTGACTATGCTGATGTGAAGTACCTAAGCCATCAGA 708
Db 15442 CACCCGGCC-----GATGTTTGAACCTCTTTAACAGTGAACAAATCCCGAGCTTTTA-- 15494

Qy 709 CTCCTACCTTTAAATATGATGTGGCGGAGGACCGCTGCTGCTGATCTGATTTCCAGC 768
Db 15495 -----TAAATATATAGCTGGGCTGGCGCGGTGCTCACACTGTAATCCCGCAGC 15545

Qy 769 ACTTTGGAGGACAGAGTGGGTGAATCACTTGAGGCCAGGAGTTTGAGACCAGCTGGCC 828
Db 15546 AGTTTGGAGGACCAAGGACGAGTGGTACCTTGAGCTCGGGTGTTCAGAGACCAGCTTGGCC 15605

Qy 829 AACATGGTGAACCTCTCTTTACTATAAATAAATAAATAAATAAATAAATAAATAAATAA 888
Db 15606 AACATGGTGAACCTCTCTTTACT-----AAAAATACAAAAATTAGCCAGGT 15653

Qy 889 GTGCTGGGACACACTCTGATCCAGCTATGCTGGAGGCTGAGGACGAGTCACTTGA 948
Db 15654 GTGCTGGGACACACTCTGATCCAGCTACTTGGAGGCTGAGGACGAGTCACTTGA 15713

Qy 949 ACCCTGGAGGCGGAGGTGGAGTGGGCGGAGATCACATCACCGCCCTCCAGCTGGCGGA 1008
Db 15714 ACCCGGAGGCGGAGGTGGAGTGGGCGGAGATTTATCCACTGCACTCCACTCTGGGTGA 15773

Qy 1009 CAGAGCAAGACTCTGTCTCAATTAATAATAAATAAATAAATAAATAAATAAATAAAGT 1058
Db 15774 AAGAGCAAGACTCTGTCTCAAAACATTAATAAATAAATAAATAAATAAATAAAGT 15823
```

## RESULT 6

AAS36444  
ID AAS36444 standard; DNA; 19345 BP.

XX AC AAS36444;

XX 17-DEC-2001 (first entry)

DE Human cardiovascular system antigen genomic DNA SEQ ID No 1944.

XX Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;  
KW chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;  
KW antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
KW cerebroprotective; nootropic; antibacterial; virucide; fungicide; cancer;  
KW ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasm;  
KW hyperproliferative disorder; breast; liver; cardiovascular disorder; ds;  
KW cerebrovascular disorder; nervous system disorder; bacterial infection;  
KW fungal infection; viral infection; ocular disorder; endocrine disorder;  
KW gastrointestinal disorder; renal disorder; respiratory disorder;  
KW wound healing; skin aging; organ transplantation; tissue regeneration;  
KW anti-infertility.  
XX Homo sapiens.  
XX WO200155321-A2.  
PN 02-AUG-2001.  
PD 17-JAN-2001; 2001WO-US01340.  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.



```
Db 15450 CACCCGGCC-----GATGTTTGCACCTCTTTAAACAGTGAACAAATCCCGAGCTTTTA-- 15502
QY 709 CTCCTACCTTTAAATATGCGAGTTTGGCCAGGACCGTGGCTCATGCTGTAAATCCAGC 768
Db 15503 -----TAAATATATAAGCTGGGCTGGGCGGTGGCTCACACTGTAAATCCAGC 15553
QY 769 ACTTTGGAGGACAGAGTGGGTGAATCACCTTGAGGCCAGGAGTTTGAGACCAGCCTGGCC 828
Db 15554 AGTTTGGAGGACCAAGGACAGTGCATCAGCTCGGGTGTTCAGACCAGCTTGGCC 15613
QY 829 AACATGGTGAACCTCTGCTTTACTATAAAAAAATAAAAAAATACAGCCGGT 888
Db 15614 AACATGGTGAACCTCTGCTTTACT-----AAAAATACAAAAATTAGCCAGT 15661
QY 889 GTCGTGGGGCACACCTGTAAATCCAGCTATGCTGGAGGCTGAGGCAGAGTCACTTGA 948
Db 15662 GTGGTGGCACACACCTGTAAATCCAGCTACTTGGGAGGCTGAAGCAGAGAACTCACTGA 15721
QY 949 ACCCTGGAGGGAGGTTGTCAGTGGGCGGAGATCACATCACCGCCCTCCAGCCTGGCGGA 1008
Db 15722 ACCCGGAGGACAGAGTTGTCAGTGAGCGCGAGATTATGCCACTGCACCTCCACTCTGGGTGA 15781
QY 1009 CAGAGCAAGACTCTGCTCAATAAATAAATAAACAACGACGAAGCAGT 1058
Db 15782 AAGAGCAAGACTCTGCTCAAAACATATAAAAAATAAATAAATAAAGT 15831

RESULT 7
AAL05001/c
ID -AAL05001 standard; DNA; 12970 BP.
XX
AC AAL05001;
XX
DT 21-NOV-2001 (first entry)
XX
DE Human reproductive system related antigen DNA SEQ ID NO: 7689.
XX
KW Human; reproductive system related antigen; reproductive system disorder;
KW cancer; gene therapy; ds.
XX
OS Homo sapiens.
XX
PN WO200155320-A2.
XX
PD
XX
PF 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US01339.
XX
PR 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 07-JUN-2000; 2000US-0205151.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 11-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226686.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0231968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0232398.
PR 14-SEP-2000; 2000US-0232399.
PR 14-SEP-2000; 2000US-0232400.
PR 14-SEP-2000; 2000US-0232401.
PR 14-SEP-2000; 2000US-0232403.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 25-SEP-2000; 2000US-0234999.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239935.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241809.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
```



PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-MAR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216980.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.

PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 08-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX XX  
(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483232/52.

Nucleic acids encoding 973 human testicular antigen polypeptides,  
useful for preventing, diagnosing and/or treating testicular cancer  
Disclosure; SEQ ID NO 2546; 766pp; English.

The present invention provides the protein and coding sequences of 973

CC human testicular antigens, and fragments of their genomic sequences. The  
CC sequences can be used in the treatment of cardiovascular, urinary system,  
CC reproductive system, immune, respiratory, neurological and  
CC gastrointestinal disorders, infections, and particularly cancer.  
CC especially testicular cancers. The present sequence is a DNA encoding a  
CC protein fragment of the invention.

XX  
SQ Sequence 12970 BP; 3310 A; 2917 C; 2938 G; 3805 T; 0 other;  
Query Match 20.7%; Score 414.8; DB 23; Length 12970;  
Best Local Similarity 73.3%; Pred No. 5.8e-64;  
Matches 608; Conservative 0; Mismatches 197; Indels 25; Gaps 5;

QY 358 TTTTCTTTTCTTTTGTGAGCGGAGTCTCCTCTGTGCGCCAGGCTGGAGTGCAGTGG 417  
DB 6179 TTTTCTTTTCTTTTGTGAGCGGAGTCTCCTCTGTGCGCCAGGCTGGAGTGCAGTGG 6120  
QY 418 CGGATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCACGCCATTCCTCGCTCAGC 477  
DB 6119 CGTATCTCGGCTCACTGCAAGCTCCGCTCCGGGTTCACGCCATTCCTCGCTCAGC 6060  
QY 478 CTCCTCAAGTAGCTGGGACTACAGCGCGCGCCACTACGCCGCTAAATTTTCTTATT 537  
DB 6059 CTCCTCAAGTAGCTGGGACTACAGCGCGCGCCACTACGCCGCTAAATTTTCTTATT 6002  
QY 538 TTAGTAGAGCGGGTTCACCGTTTGTAGCGGGATGGCTCGATCTCTCGACCTCTGCA 597  
DB 6001 TTAGTAGACACAGGCTTTCACCGTGTAGCCAGGATGGTCTCATCTCTCGACCTCTGA 5942  
QY 598 TCCGCCCGGCTCGGCTCCCAAGTGTGGGATTACAGCGGTGAGCCACCGCGCCCGCC 657  
DB 5941 TCTGCTCGGCTCGGCTCCCAAGTGTGGGATTACAGCGGTGAGCCACCGCGCCCG 5882  
QY 658 ATGA-TCACTCTCTTGACTATCTGATGTGCAAGTACTTAAGCCATCAGACTCTACCC 716  
DB 5881 ACCACTAGTGTTTTAATGGGTAAATAATTTGTACAGATACCTTAACAAC----- 5829  
QY 717 TTTAATATGCAAGTTTGGCCAGGCGGCTGATGCTGCTGTAATTCAGACCTTTGGG 776  
DB 5828 -GAAATATGTGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5770  
QY 777 AGGAGAGTGGTGAATCACTTGTAGGCGAGAGTTGAGACCGCTGGCGCAACATGGT 836  
DB 5769 AGGCGGAGTGGGCGGAGATCAC--GAGGTCAAGAGATCGAGACCATCTGGCTTAACAC 5712  
QY 837 GAACTCTGCTTTTACTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 896  
DB 5711 GAAACCTGCTCTCTAAATAAT-----ATAAATAATAGCCATGCGTGGTGAC 5664  
QY 897 GCACACCTGTAATCCAGCTAGTGTGGGCGTGGGCGGCGGCGGCGGCGGCGGCGGCGG 956  
DB 5663 AGGCACCTGTAGTCCAGCTACTTGGGAGGCTGATGAGGAGATTTGCTCAACCGGGA 5604  
QY 957 GCGGAGTGTGAGTGGCGGAGATCAATCAACCGCTTCCAGCTGGCGGAGAGAGAA 1016  
DB 5603 GGTGAGGCTGCGTGTGAGGAGATCAATCAACCGCTTCCAGCTGGCGGAGAGAGAA 5544  
QY 1017 GACTCTGCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 1076  
DB 5543 GACTCCATCTCAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 5484  
QY 1077 ATCTAAAAAATGCTGTCAACAAATAGAGCAAGTGAATAAAGGAAATAAATGG 1136  
DB 5483 ATGCACATGAAGAAGCTTTAGTCAATCAGAAAAATGGAATTAAGACACATGAATA 5424  
QY 1137 GCCAAGACTCTPAAGTATATTGTCAAAATCATTCAGAACCTTTAAAAA 1186  
DB 5423 CCAATACACACCCAGTAGAATGGCTAAATAAATAAAGAACACTGACAAACAA 5374

RESULT 9  
ABK94412/c  
ID ABK94412 standard; DNA; 74037 BP.

XX  
AC ABK94412;  
XX  
DT 27-AUG-2002 (first entry)  
XX  
DE DNA encoding endothelin converting enzyme 1 (ECE-1) #2.  
XX  
KW Endothelin; EDN; endothelin converting enzyme 1; ECE 1; EDNR;  
KW endothelin receptor; signaling system; cardiovascular disease;  
KW coronary heart disease; hypertension; atherosclerosis; angiogenesis;  
KW fatty acid metabolism; diabetes; familial hypercholesterolaemia;  
KW forensic marker; transgenic animal; solid support; SNP;  
KW cardiovascular regulator; gene; ds; single nucleotide polymorphism.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT variation replace(8625,T)  
FT /\*tag= a  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(8938,T)  
FT /\*tag= b  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(9351,G)  
FT /\*tag= c  
FT /\*standard\_name= "Single nucleotide polymorphism"  
FT variation replace(9439,C)  
FT /\*tag= d  
FT /\*standard\_name= "Single nucleotide polymorphism"  
XX  
PN WO200224747-A2.  
XX  
PD 28-MAR-2002.  
XX  
PF 31-AUG-2001; 2001WO-BP10087.  
XX  
PR 19-SEP-2000; 2000EP-0120123.  
XX (EPID-) EPIDAUROS BIOTECHNOLOGIE AG.  
XX Brinkmann U, Hoffmeyer S; ;  
XX WPI; 2002-435060/46.  
XX  
PT Novel polynucleotide of the endothelin/endothelin converting  
PT enzyme/receptors of endothelin and endothelin converting enzyme  
PT signaling system associated with cardiovascular disease, useful for  
PT treating the disease  
XX  
PS Claim 1; Page -; 190pp; English.  
XX  
CC The invention describes a polynucleotide (I) of the endothelin  
CC (EDN)/endothelin converting enzyme (ECE)/receptors of EDN and ECE (EDNR)  
CC signaling system which is associated with a cardiovascular disease. (I),  
CC the gene encoding EDN, ECE or EDNR (II) or a vector (III) expressing (I)  
CC or (II) is useful for producing cells capable of expressing a molecular  
CC variant polypeptide which is associated with a cardiovascular disease.  
CC (II), (III), the EDN, ECE or EDNR polypeptide, or a cell expressing  
CC a molecular variant gene comprising (I) is useful for identifying and  
CC obtaining a pro-drug or drug capable of modulating the activity of a  
CC molecular variant of a polypeptide of the EDN/EDNR/ECE signaling system  
CC or its gene product, or for identifying and obtaining an inhibitor of  
CC the activity of a molecular variant of a polypeptide of the EDN/EDNR/ECE  
CC signaling system or its gene product. The isolated proteins and  
CC polynucleotides encoding them are useful for preparation of a  
CC pharmaceutical composition for treating a cardiovascular disease such as  
CC coronary heart disease, hypertension, atherosclerosis, or related to  
CC abnormal angiogenesis or fatty acid metabolism e.g. diabetes and familial  
CC hypercholesterolaemia. The gene or a polynucleotide fragment of the  
CC EDN/ECE/EDNR signaling system are useful as forensic markers, for  
CC creating a transgenic animal and in creation of a solid support  
CC comprising polynucleotides, genes, vectors, polypeptides, antibodies or  
CC host cells of the invention. This sequence encodes a fragment of the



PR 27-SEP-2000: 2000US-0235836.  
PR 29-SEP-2000: 2000US-0236327.  
PR 29-SEP-2000: 2000US-0236367.  
PR 29-SEP-2000: 2000US-0236368.  
PR 29-SEP-2000: 2000US-0236369.  
PR 29-SEP-2000: 2000US-0236370.  
PR 02-OCT-2000: 2000US-0236370.  
PR 02-OCT-2000: 2000US-0236802.  
PR 02-OCT-2000: 2000US-0237037.  
PR 02-OCT-2000: 2000US-0237038.  
PR 02-OCT-2000: 2000US-0237039.  
PR 02-OCT-2000: 2000US-0237040.  
PR 13-OCT-2000: 2000US-0239935.  
PR 13-OCT-2000: 2000US-0239937.  
PR 20-OCT-2000: 2000US-0240960.  
PR 20-OCT-2000: 2000US-0241221.  
PR 20-OCT-2000: 2000US-0241785.  
PR 20-OCT-2000: 2000US-0241786.  
PR 20-OCT-2000: 2000US-0241787.  
PR 20-OCT-2000: 2000US-0241808.  
PR 20-OCT-2000: 2000US-0241809.  
PR 20-OCT-2000: 2000US-0241826.  
PR 01-NOV-2000: 2000US-024617.  
PR 08-NOV-2000: 2000US-0246474.  
PR 08-NOV-2000: 2000US-0246475.  
PR 08-NOV-2000: 2000US-0246476.  
PR 08-NOV-2000: 2000US-0246477.  
PR 08-NOV-2000: 2000US-0246478.  
PR 08-NOV-2000: 2000US-0246523.  
PR 08-NOV-2000: 2000US-0246524.  
PR 08-NOV-2000: 2000US-0246525.  
PR 08-NOV-2000: 2000US-0246526.  
PR 08-NOV-2000: 2000US-0246527.  
PR 08-NOV-2000: 2000US-0246528.  
PR 08-NOV-2000: 2000US-0246532.  
PR 08-NOV-2000: 2000US-0246609.  
PR 08-NOV-2000: 2000US-0246610.  
PR 08-NOV-2000: 2000US-0246611.  
PR 17-NOV-2000: 2000US-0246613.  
PR 17-NOV-2000: 2000US-0249207.  
PR 17-NOV-2000: 2000US-0249208.  
PR 17-NOV-2000: 2000US-0249209.  
PR 17-NOV-2000: 2000US-0249210.  
PR 17-NOV-2000: 2000US-0249211.  
PR 17-NOV-2000: 2000US-0249212.  
PR 17-NOV-2000: 2000US-0249213.  
PR 17-NOV-2000: 2000US-0249214.  
PR 17-NOV-2000: 2000US-0249215.  
PR 17-NOV-2000: 2000US-0249216.  
PR 17-NOV-2000: 2000US-0249217.  
PR 17-NOV-2000: 2000US-0249218.  
PR 17-NOV-2000: 2000US-0249244.  
PR 17-NOV-2000: 2000US-0249245.  
PR 17-NOV-2000: 2000US-0249246.  
PR 17-NOV-2000: 2000US-0249264.  
PR 17-NOV-2000: 2000US-0249265.  
PR 17-NOV-2000: 2000US-0249297.  
PR 17-NOV-2000: 2000US-0249299.  
PR 17-NOV-2000: 2000US-0249300.  
PR 01-DEC-2000: 2000US-0249300.  
PR 01-DEC-2000: 2000US-0250160.  
PR 01-DEC-2000: 2000US-0250391.  
PR 05-DEC-2000: 2000US-0251030.  
PR 05-DEC-2000: 2000US-0251988.  
PR 05-DEC-2000: 2000US-0256719.  
PR 08-DEC-2000: 2000US-0251479.  
PR 08-DEC-2000: 2000US-0251856.  
PR 08-DEC-2000: 2000US-0251868.  
PR 08-DEC-2000: 2000US-0251869.  
PR 08-DEC-2000: 2000US-0251989.  
PR 08-DEC-2000: 2000US-0251990.  
PR 11-DEC-2000: 2000US-0254097.  
PR 05-JAN-2001: 2001US-0259678.

(HUMA-) HUMAN GENOME SCI INC.

PI Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-465570/50.  
DR Isolated nucleic acid molecule encoding a reproductive system antigen  
XX PT is used in preventing, treating or ameliorating a medical condition -  
PT  
XX Disclosure; SEQ ID NO 6380; 1297pp + Sequence Listing; English.  
XX  
CC The present invention provides the protein and coding sequences of a  
CC number of human reproductive system related antigens. These can be used  
CC in the prevention and treatment of reproductive system disorders,  
CC including cancer. The present sequence is a genomic sequence encoding a  
CC protein of the invention.  
XX  
SQ Sequence 10445 BP; 2469 A; 2885 C; 2678 G; 2413 T; 0 other;  
  
Query Match 20.5%; Score 410.4; DB 22; Length 10445;  
Best Local Similarity 78.9%; Pred. No. 3.4e-63;  
Matches 560; Conservative 0; Mismatches 126; Indels 24; Gaps 5;  
  
QY 359 TTTT TTTT TTTT TTTT TGACGAGTCTCGCTGTCGCCAGGCTGGAGTCAGTGGC 418  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 8627 TTTT TTTT TTTT TTTT TGACGAGTCTCGCTGTCGCCAGGCTGGAGTCAGTGGC 8686  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
QY 419 GGGATCTGGCTCACTGCAAGCTCCGCTCCCGGTTTCAGGCCATTCTCTGCTCAGCC 478  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 8687 GCTATCTGGCTCACTGCAAGCTCTGCTCCCGGTTTCAGGCCATTCTCTGCTCAGTC 8746  
QY 479 TCCCAAGTAGCTGGGACTACAGGCGCCGCTACGCTACGCGCGGCTAAATTTTGTATTTT 538  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 8747 TCCCAAGTAGCTGGGACTACAGGCGCCGCTACGCTACGCGCGGCTAAATTTTGTATTTT 8805  
QY 539 TAGTAGAGACGGGTTTTCACCGTTTAGCCGGGATGGCTCGATCTCTGACCTCGTGAT 598  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 8806 TAGTAGAGACGGGTTTTCACCATGTTAGCCAGGATGTCGATCTCTGACCTCATGAT 8865  
QY 599 CCGCCGCTCGGCTCCCAAGTGTGGATTACAGCGGTACAGCCAGCCGCGCGGCCA 658  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 8866 CCACCGCTCAGCTCCCAAGTGTGGATTACAGCGGTACAGCCAGCCGCGCGGCC 8925  
QY 659 TGATCATCTTCTTACTATGCTGATGTGACAAGTACCTAAGCCATCAGACTCTACCCCTT 718  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 8926 AGAATCTGCATTTTAAAAAATCCCAAGTCTTAAGTACAGGGGATCTTAG----- 8976  
QY 719 TAAATATGCACTTTGGCCAGGACCGTGTGCTCATGCTGTAATTCAGCAGCTTTGGAG 778  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 8977 AATAAGAAATTTAGGCTGGGCTGGTGTGCTCAAGCTGTAATCTCAGCA-TTTGGGAG 9035  
QY 779 GCAGAGTGGTGAATCACTTGAGGCCAGGAGTTGAGACAGCTTGCCCAACATGTTGA 838  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 9036 GCGGGGTGGCAGATTACCTGGGTCAGGAGTTCAAGACAGCCAGCTGTCACATGTTGA 9095  
QY 839 AACTCTGTCTTTACTAAAAAATAAAAAAATAAAAAAATAACAGCCGGGTGCTGGGGC 898  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 9096 AACCCCATCTGTACT-----AAAAACAATAAAAAATTTGCTGCCATCGTGTGG 9144  
QY 899 ACACCTGTATCCAGCTATGCTGGAGGCTGAGGCACGAGAGTCACTTGAACCTTGGAGG 958  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 9145 GTACCTTTAATCCAGCTACTCAGGAGGCTGAGCAGGAGAAATCGTTGAACCTTGGAGG 9204  
QY 959 CGGAGGTTGCAGTGGGCGGAGATCAGATCAGCCGCTCCAGCCCTGGGCGAC--AGAGCAA 1016  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 9205 TGGAGGTTGAGTGAGCAAGATCATGCTACTGCATCCAGCTGGCGGCAAGAGTGA 9264  
QY 1017 GACTCTGTCTCAATAATAATAACAAACAGCAAGACAGTGTGTTGTA 1066  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 9265 AACTCCATCTCAAAAAAATAAAAAAATAAAAAAATAAGAAAGAAATGTTTGA 9314  
  
RESULT 11  
AAL05001  
ID AAL05001 standard; DNA; 12970 BP.

XX AAL05001; 14-SEP-2000; 2000US-0232400.  
AC 14-SEP-2000; 2000US-0232401.  
XX 14-SEP-2000; 2000US-0233053.  
DT 14-SEP-2000; 2000US-0233054.  
XX 14-SEP-2000; 2000US-0233065.  
XX 21-SEP-2000; 2000US-0234223.  
DE 21-SEP-2000; 2000US-0234274.  
XX 25-SEP-2000; 2000US-0234997.  
KW 25-SEP-2000; 2000US-0234998.  
KW 26-SEP-2000; 2000US-0235484.  
XX 27-SEP-2000; 2000US-0235834.  
OS 27-SEP-2000; 2000US-0235836.  
XX 29-SEP-2000; 2000US-0236327.  
PN 29-SEP-2000; 2000US-0236367.  
XX 29-SEP-2000; 2000US-0236368.  
XX 29-SEP-2000; 2000US-0236369.  
XX 29-SEP-2000; 2000US-0236370.  
PD 29-SEP-2000; 2000US-0236370.  
XX 02-OCT-2000; 2000US-0236802.  
PF 02-OCT-2000; 2000US-0237037.  
XX 02-OCT-2000; 2000US-0237038.  
XX 02-OCT-2000; 2000US-0237039.  
XX 02-OCT-2000; 2000US-0237040.  
XX 13-OCT-2000; 2000US-0239935.  
XX 13-OCT-2000; 2000US-0239937.  
XX 20-OCT-2000; 2000US-0240960.  
XX 20-OCT-2000; 2000US-0241221.  
XX 20-OCT-2000; 2000US-0241785.  
XX 20-OCT-2000; 2000US-0241786.  
XX 20-OCT-2000; 2000US-0241787.  
XX 20-OCT-2000; 2000US-0241808.  
XX 20-OCT-2000; 2000US-0241809.  
XX 20-OCT-2000; 2000US-0241826.  
XX 01-NOV-2000; 2000US-0244617.  
XX 08-NOV-2000; 2000US-0246474.  
XX 08-NOV-2000; 2000US-0246475.  
XX 08-NOV-2000; 2000US-0246476.  
XX 08-NOV-2000; 2000US-0246477.  
XX 08-NOV-2000; 2000US-0246478.  
XX 08-NOV-2000; 2000US-0246523.  
XX 08-NOV-2000; 2000US-0246524.  
XX 08-NOV-2000; 2000US-0246525.  
XX 08-NOV-2000; 2000US-0246526.  
XX 08-NOV-2000; 2000US-0246527.  
XX 08-NOV-2000; 2000US-0246528.  
XX 08-NOV-2000; 2000US-0246532.  
XX 08-NOV-2000; 2000US-0246609.  
XX 08-NOV-2000; 2000US-0246610.  
XX 08-NOV-2000; 2000US-0246611.  
XX 08-NOV-2000; 2000US-0246613.  
XX 17-NOV-2000; 2000US-0249207.  
XX 17-NOV-2000; 2000US-0249208.  
XX 17-NOV-2000; 2000US-0249209.  
XX 17-NOV-2000; 2000US-0249210.  
XX 17-NOV-2000; 2000US-0249211.  
XX 17-NOV-2000; 2000US-0249212.  
XX 17-NOV-2000; 2000US-0249213.  
XX 17-NOV-2000; 2000US-0249214.  
XX 17-NOV-2000; 2000US-0249215.  
XX 17-NOV-2000; 2000US-0249216.  
XX 17-NOV-2000; 2000US-0249217.  
XX 17-NOV-2000; 2000US-0249218.  
XX 17-NOV-2000; 2000US-0249244.  
XX 17-NOV-2000; 2000US-0249245.  
XX 17-NOV-2000; 2000US-0249245.  
XX 17-NOV-2000; 2000US-0249264.  
XX 17-NOV-2000; 2000US-0249265.  
XX 17-NOV-2000; 2000US-0249297.  
XX 17-NOV-2000; 2000US-0249299.  
XX 17-NOV-2000; 2000US-0249300.  
XX 01-DEC-2000; 2000US-0250160.  
XX 01-DEC-2000; 2000US-0250391.  
XX 05-DEC-2000; 2000US-0251030.  
XX 05-DEC-2000; 2000US-0251988.  
XX 05-DEC-2000; 2000US-0256719.  
XX 05-DEC-2000; 2000US-0256719.  
XX 31-JAN-2000; 2000US-0179065.  
XX 04-FEB-2000; 2000US-0180628.  
XX 24-FEB-2000; 2000US-0184664.  
XX 02-MAR-2000; 2000US-0186350.  
XX 16-MAR-2000; 2000US-0189874.  
XX 17-MAR-2000; 2000US-0190076.  
XX 18-MAR-2000; 2000US-0198123.  
XX 19-MAY-2000; 2000US-0205515.  
XX 07-JUN-2000; 2000US-0209467.  
XX 28-JUN-2000; 2000US-0214886.  
XX 30-JUN-2000; 2000US-0215135.  
XX 07-JUL-2000; 2000US-0216647.  
XX 07-JUL-2000; 2000US-0216880.  
XX 11-JUL-2000; 2000US-0217487.  
XX 11-JUL-2000; 2000US-0217496.  
XX 14-JUL-2000; 2000US-0218290.  
XX 26-JUL-2000; 2000US-0220963.  
XX 26-JUL-2000; 2000US-0220964.  
XX 14-AUG-2000; 2000US-0224518.  
XX 14-AUG-2000; 2000US-0224519.  
XX 14-AUG-2000; 2000US-0225213.  
XX 14-AUG-2000; 2000US-0225214.  
XX 14-AUG-2000; 2000US-0225256.  
XX 14-AUG-2000; 2000US-0225267.  
XX 14-AUG-2000; 2000US-0225268.  
XX 14-AUG-2000; 2000US-0225270.  
XX 14-AUG-2000; 2000US-0225278.  
XX 14-AUG-2000; 2000US-0225447.  
XX 14-AUG-2000; 2000US-0225757.  
XX 14-AUG-2000; 2000US-0225758.  
XX 14-AUG-2000; 2000US-0225759.  
XX 18-AUG-2000; 2000US-0226279.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0226681.  
XX 22-AUG-2000; 2000US-0227182.  
XX 23-AUG-2000; 2000US-0227009.  
XX 30-AUG-2000; 2000US-0228924.  
XX 01-SEP-2000; 2000US-0229287.  
XX 01-SEP-2000; 2000US-0229343.  
XX 01-SEP-2000; 2000US-0229344.  
XX 01-SEP-2000; 2000US-0229345.  
XX 05-SEP-2000; 2000US-0229509.  
XX 05-SEP-2000; 2000US-0229513.  
XX 06-SEP-2000; 2000US-0230437.  
XX 06-SEP-2000; 2000US-0230438.  
XX 08-SEP-2000; 2000US-0231242.  
XX 08-SEP-2000; 2000US-0231243.  
XX 08-SEP-2000; 2000US-0231244.  
XX 08-SEP-2000; 2000US-0231413.  
XX 08-SEP-2000; 2000US-0231414.  
XX 08-SEP-2000; 2000US-0232080.  
XX 08-SEP-2000; 2000US-0232081.  
XX 12-SEP-2000; 2000US-0231968.  
XX 14-SEP-2000; 2000US-0232397.  
XX 14-SEP-2000; 2000US-0232398.  
XX 14-SEP-2000; 2000US-0232399.  
XX 14-SEP-2000; 2000US-0232399.  
Human reproductive system related antigen DNA SEQ ID NO: 7689.  
Human; reproductive system related antigen; reproductive system disorder;  
cancer; gene therapy; ds.  
Homo sapiens.  
W0200155320-A2.  
02-AUG-2001.  
17-JAN-2001; 2001WO-US01339.



	PR	17-NOV-2000;	2000US-0249214.	
	PR	17-NOV-2000;	2000US-0249215.	
	PR	17-NOV-2000;	2000US-0249216.	
	PR	17-NOV-2000;	2000US-0249217.	
	PR	17-NOV-2000;	2000US-0249218.	
	PR	17-NOV-2000;	2000US-0249219.	
	PR	17-NOV-2000;	2000US-0249220.	
	PR	17-NOV-2000;	2000US-0249221.	
	PR	17-NOV-2000;	2000US-0249222.	
	PR	17-NOV-2000;	2000US-0249223.	
	PR	17-NOV-2000;	2000US-0249224.	
	PR	17-NOV-2000;	2000US-0249225.	
	PR	17-NOV-2000;	2000US-0249226.	
	PR	17-NOV-2000;	2000US-0249227.	
	PR	17-NOV-2000;	2000US-0249228.	
	PR	17-NOV-2000;	2000US-0249229.	
	PR	17-NOV-2000;	2000US-0249230.	
	PR	01-DEC-2000;	2000US-0250160.	
	PR	01-DEC-2000;	2000US-0250391.	
	PR	05-DIC-2000;	2000US-0251030.	
	PR	05-DIC-2000;	2000US-0251988.	
	PR	05-DIC-2000;	2000US-0256719.	
	PR	06-DEC-2000;	2000US-0251479.	
	PR	08-DEC-2000;	2000US-0251856.	
	PR	08-DEC-2000;	2000US-0251868.	
	PR	08-DEC-2000;	2000US-0251869.	
	PR	08-DEC-2000;	2000US-0251989.	
	PR	08-DEC-2000;	2000US-0251990.	
	PR	11-DEC-2000;	2000US-0254097.	
	PR	05-JAN-2001;	2000US-0259678.	
	XX	(HUMA-) HUMAN GENOME SCI INC.		
	PA	Rosen CA,	Barash SC,	Ruben SM;
	PI	WPI; 2001-483232/52.		
	DR	Nucleic acids encoding 973 human testicular antigen polypeptides, useful for preventing., diagnosing and/or treating testicular cancer -		
	XX	Disclosure; SEQ ID NO 2546; 766pp; English.		
	PS	The present invention provides the protein and coding sequences of 973 human testicular antigens, and fragments of their genomic sequences. The sequences can be used in the treatment of cardiovascular, urinary system, reproductive system, immune, respiratory, neurological and gastrointestinal disorders, infections, and particularly cancer, especially testicular cancers. The present sequence is a DNA encoding a protein fragment of the invention.		
	CC	Sequence 12970 BP; 3310 A; 2917 C; 2938 G; 3805 T; 0 other;  Query Match            20.3%; Score 406.2; DB 23; Length 12970; Best Local Similarity 75.7%; Pred. No. 1.9e-62; Matches 564; Conservative 0; Mismatches 158; Indels 23; Gaps 4;		
	QY	342 ATCTTCCTTTTTTTTGTGAGACGGGATCTCGCCTCTCGTGCGGCCA 401		
	DB	5499 ATTTCCTTTGGTTTTTGGTTTTTTTGTGAGATTGCCTCTCGCCCCA 5558		
	QY	402 GGCTGGAGTCACTGCGGGAATCTCGCTCACAGAAGCTCCGCCCTCCGGGTTCACGCC 461		
	DB	5559 GGCTGGAGTCACTGATATGATCTGGCTCATCTGCAGCTCCCGGGTTGGAGCA 5618		
	QY	462 ATTCTCCTCGCTCAGCCTCCCAAAGTAGTGGGACTACAGCGCCCGCCACTACGCCCGGC 521		
	DB	5619 ATTCTCCTGCATCAGCCTCCCAAAGTAGTGGGACTACAGGTGCTGTACCACCGCATGCC 5678		
	QY	522 TAATTTTTTTGTATTTTAGTAGAGAGGGGTTTACCGTTTTAGCGGGGATGGCCTCGA 581		
	DB	5679 TAA--TTTTTATATTTTAGTAGAGAGAGGGTTTACCGGTGTTAGCCAGGATGGTCTCGA 5736		
	QY	582 TCCTCCGACCTCGTGATCGCCCGCTCGSCCTCCCAAAGTGTCTGGGATTACAGGCGTGA 641		
	DB	5737 TCFTTTGACCTCGTGATCTGCCCACTCGGCCCTCCCAAAGTGTCTGGGATTACAGGCGTGA 5796		
	QY	642 GCCACGCGCGCCGCGCATGATCTTCTTTGACTATCTGATGTGACAAGTAACCTAAAGC 701		

Db	5797	GCCACCGCGCTGGCCGCCGCACATATTTTCGTTTGAAGTATCTGTACAAATATTTT	5856	PR	14-AUG-2000;	2000US-0225211.30
Qy	702	CATCAGACTCTACCCCTTTAAATATGTCAGTTTGGCCAGGCACCGTGCCTCATGCTCTAA	761	PR	14-AUG-2000;	2000US-0225214
Db	5857	TA-----CCATTTAAAAACACTGAGTGTGTCGGGCACAGTGGCTCACGCTGTAA	5909	PR	14-AUG-2000;	2000US-0225267
Qy	762	TTCCAGCACTTTGGGAGCGAGAGTGGGTGAATCACTTGAGCCAGGAGTTTGAGACCAG	821	PR	14-AUG-2000;	2000US-0225268
Db	5910	TCCCAGAACTTTGGGAGCGCGAGCGAGCAGATCAA--GAGGTGAGGAGATGGAGACCAT	5967	PR	14-AUG-2000;	2000US-0225270
Qy	822	CCTGGCCAACTGTGTAAACTCTGTCTTTACTAAAAAATAAATAAATAAATAAATAA	881	PR	14-AUG-2000;	2000US-0225271
Db	5968	CCTGGCTAACCGTGAACCCCTGTGTCTACT-----AAAAATAAATAAATAAATAA	6015	PR	14-AUG-2000;	2000US-0225272
Qy	882	GCCGGGTGTCTGGGGCACACCTCTAATCCAGCTATGCTGGAGGCTGAGGCAGCAGAGT	941	PR	14-AUG-2000;	2000US-0225273
Db	6016	GCCGGGATGTTGGCAGGCGCCTGTAGTCCAGCTACTCGGGAGGCTGAGGCAGGAGAT	6075	PR	14-AUG-2000;	2000US-0225274
Qy	942	CACCTTGAACCTGGAGGCGGAGGTTGCAGTGGCGCGGAGATCACATCACCGCCCTCCAGCC	1001	PR	14-AUG-2000;	2000US-0225275
Db	6076	GGCGTGAACCCAGGAAGTGGAGCTTGCAGTGAGCGCGGAGATCACGCCACTGCCTCCAGCC	6135	PR	14-AUG-2000;	2000US-0225276
Qy	1002	TGGCGGCACAGACGAAGCACTCTGTCTCAATAAATAAATAAATAAATAAATAAATAA	1061	PR	14-AUG-2000;	2000US-0225277
Db	6136	TGGCGCAGACTGCAGACTCCGCTCTCAAAAAAATAAATAAATAAATAAATAAATAA	6195	PR	14-AUG-2000;	2000US-0225278
Qy	1062	TTGTACCTTAGTTTATATCTAAAAA	1086	PR	14-AUG-2000;	2000US-0225279
Db	6196	TTGAGTTGTAGGAATGCTTATATA	6220	PR	14-AUG-2000;	2000US-0225280
RESULT 13						
AAK89986/C						
ID	AAK89986 standard; DNA; 32224 BP.					
XX	AAK89986;					
AC						
DT	05-NOV-2001 (first entry)					
DE	Human digestive system antigen genomic sequence SEQ ID NO: 3562.					
DE	Human; digestive system antigen; gene therapy; cancer; appendicitis;					
KW	ulcerative colitis; Infection; Hirschsprung's disease; chronic colitis;					
KW	digestive system disorder; Meckel's diverticulum; ds.					
KW						
OS	Homo sapiens.					
OS						
XX	WO200155314-A2.					
PN	02-AUG-2001.					
PN	17-JAN-2001; 2001WO-US01324.					
PD						
PD						
PF						
PF						
XX						
XX						
PR	31-JAN-2000; 2000US-0179065.					
PR	04-FEB-2000; 2000US-0180628.					
PR	24-FEB-2000; 2000US-0184664.					
PR	02-MAR-2000; 2000US-0186350.					
PR	16-MAR-2000; 2000US-0189874.					
PR	17-MAR-2000; 2000US-0190076.					
PR	18-APR-2000; 2000US-0198123.					
PR	19-MAY-2000; 2000US-0205515.					
PR	07-JUN-2000; 2000US-0209467.					
PR	28-JUN-2000; 2000US-0214886.					
PR	30-JUN-2000; 2000US-0215135.					
PR	07-JUL-2000; 2000US-0216647.					
PR	07-JUL-2000; 2000US-0216880.					
PR	11-JUL-2000; 2000US-0217487.					
PR	11-JUL-2000; 2000US-0217496.					
PR	14-JUL-2000; 2000US-0218290.					
PR	26-JUL-2000; 2000US-0220963.					
PR	26-JUL-2000; 2000US-0220964.					
PR	14-AUG-2000; 2000US-0224518.					
PR	14-AUG-2000; 2000US-0224519.					



PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226688.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.  
PR 29-SEP-2000; 2000US-0236327.  
PR 29-SEP-2000; 2000US-0236367.  
PR 29-SEP-2000; 2000US-0236368.  
PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 20-OCT-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0242221.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250391.  
PR 01-DEC-2000; 2000US-0251160.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Barash SC, Ruben SM;  
XX WPI; 2001-541565/60.  
XX Nucleic acids encoding 3224 human nervous system antigen polypeptides,  
PT useful for preventing, diagnosing and/or treating nervous system  
PT cancers and metastases -  
XX



PR	17-NOV-2000;	2000US-024924	
PR	17-NOV-2000;	2000US-024926	
PR	17-NOV-2000;	2000US-024926	
PR	17-NOV-2000;	2000US-024926	
PR	17-NOV-2000;	2000US-024927	
PR	17-NOV-2000;	2000US-024929	
PR	17-NOV-2000;	2000US-024930	
PR	01-DEC-2000;	2000US-0250160	
PR	01-DEC-2000;	2000US-0250391	
PR	05-DEC-2000;	2000US-0251030	
PR	05-DEC-2000;	2000US-0251988	
PR	05-DEC-2000;	2000US-0256719	
PR	06-DEC-2000;	2000US-0251479	
PR	08-DEC-2000;	2000US-0251856	
PR	08-DEC-2000;	2000US-0251868	
PR	08-DEC-2000;	2000US-0251869	
PR	08-DEC-2000;	2000US-0251989	
PR	08-DEC-2000;	2000US-0251990	
PR	11-DEC-2000;	2000US-0254097	
PR	05-JAN-2001;	2001US-02595678	
XX			
PA	(HUMA-)	HUMAN GENOME SCI INC.	
XX			
PI	Rosen CA, Barash SC, Ruben SM;		
XX			
DR	WPI; 2001-451929/48.		
XX			
PT	Isolated polypeptide for treating, preventing and/or prognosing		
PT	disorders related to the reproductive system including prostate cancer		
PT	and also for testing and detection e.g. diagnosis -		
XX			
PS	Disclosure; SEQ ID NO 668; 546pp; English.		
XX			
CC	The invention relates to novel isolated human prostate cancer antigen		
CC	polynucleotides (I) and polypeptides (II). (I) and (II) are useful for		
CC	preventing, treating or ameliorating a medical condition when		
CC	administered. (I), (II) and the antibody to (II) are useful for treating,		
CC	preventing and/or prognosing disorders related to the reproductive		
CC	system including prostate cancers; urinary disorders e.g. chronic		
CC	nephritis; and blood-related disorders e.g. thrombosis. (II) can be used		
CC	for testing and detection e.g. as a chromosomal marker and in forensics.		
CC	(I) and the anti-(II) antibody can be used in testing and detection in		
CC	immunoassays. AAS40061-AAS40775 represent the human prostate cancer		
CC	antigen coding sequences, and related PCR primers and sequences of the		
CC	invention. Note: The sequence data for this patent did not form		
CC	part of the printed specification, but was obtained in electronic format		
CC	directly from WIPO at: ftp.wipo.int/pub/published_pct_sequences.		
XX			
SQ	Sequence 4433 BP; 1279 A; 897 C; 857 G; 1400 T; 0 other;		
	Query Match	20.3%; Score 405.8; DB 22; Length 4433;	
	Best Local Similarity	74.8%; Pred. No. 2e-62;	
	Matches	557; Conservative 0; Mismatches 167; Indels 21; Gaps 3;	
QY	309	AGAAGCACAGTTTCAGGGTGGTGCTATGATCATCTCTTTTTTTTTTTTTTTTTTTT 368	
Db	3741	AGGAGGAAAAGTCAAAGAAATATGACTAGCAATACTAAATCTTTTTTTTTTTTTTT 3682	
QY	369	TTTTTTTGGACGGAGTCTCGCTCTGTGCGCCAGGCTGGAGTGGCGGGATCTCGG 428	
Db	3681	TTTTTTTGGACGGAGTCTTGTCTGTGTCACCCAGGCTGGAGTGGACGATCTTGG 3622	
QY	429	CTCACTGCAAGTCGCGCTCCCGGGTTACGCGCAATCTCTGCTCAGCTCCCAAGTAG 488	
Db	3621	CTCACTGCAAGTCTTGCTCTCCAGGTTTACACCAATCTCTCTGCTTACGCTCCCAAGTAG 3562	
QY	489	CTGGGACTACAGCGCGCGGCTACGCGCGGGCTAAATTTTTTTTGTATTTTAGTAGAGAC 548	
Db	3561	CTGGGACTACAGCGGATGCCACACGCTGGCTAAATTTTTTTGTATTTTAGTAGAGAC 3502	
QY	549	GGGGTTTACCGTTTATAGCGGGATGCCCTCGATCTCCTGACCTCGTGATCCGCCGCGCT 608	
Db	3501	AGGATTTTACCATGTTAGCCAGGATGCTCTCCATCTCCTGACTTGTGTGATCCGCCGCGCT 3442	



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 5, 2003, 02:24:45 ; Search time 3765.66 Seconds  
(without alignments)  
12908.483 Million cell updates/sec

Title: US-09-936-271B-13\_COPY\_1\_2000

Perfect score: 2000

Sequence: 1 gggccagagtgaagcaag.....gtgcgtctgcaccacatc 2000

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*\*

2: em\_esthum:\*\*

3: em\_estin:\*\*

4: em\_estmu:\*\*

5: em\_estov:\*\*

6: em\_estpl:\*\*

7: em\_estro:\*\*

8: em\_htc:\*\*

9: gb\_est1:\*\*

10: gb\_est2:\*\*

11: gb\_htc:\*\*

12: gb\_est3:\*\*

13: gb\_est4:\*\*

14: gb\_est5:\*\*

15: em\_estfun:\*\*

16: em\_estom:\*\*

17: em\_gss\_hum:\*\*

18: em\_gss\_inv:\*\*

19: em\_gss\_pln:\*\*

20: em\_gss\_vrt:\*\*

21: em\_gss\_fun:\*\*

22: em\_gss\_nam:\*\*

23: em\_gss\_mus:\*\*

24: em\_gss\_pro:\*\*

25: em\_gss\_rod:\*\*

26: em\_gss\_phg:\*\*

27: em\_gss\_vrl:\*\*

28: gb\_gssl:\*\*

29: gb\_gssl2:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	397.6	19.9	2371	28	AF101960
2	373.8	18.7	2971	28	AF101960
3	352.2	17.6	929	9	AL580585
C 4	337.4	16.9	652	9	AU120416

C 5	337.2	16.9	1641	11	BC033224	BC033224 Homo sapi
C 6	336	16.8	2772	11	BC038630	BC038630 Homo sapi
C 7	334.4	16.7	698	12	BM99040	BM99040 UI-H-DIO-
8	332	16.6	969	13	BX415714	BX415714 BX415714
C 9	329.6	16.5	929	9	AL580585	AL580585 AL580585
C 10	325.2	16.3	652	9	AU120416	AU120416 AU120416
C 11	324	16.2	701	10	BE744242	BE744242 601576428
C 12	321.6	16.1	2772	11	BC038630	BC038630 Homo sapi
C 13	316.8	15.8	381	29	CC061638	CC061638 MUGQ_CH25
C 14	316.6	15.8	709	29	AG010128	AG010128 Homo sapi
15	315	15.8	710	14	CA428305	CA428305 UI-H-DFO-
16	312.6	15.6	542	9	AI679782	AI679782 tu7610.x
17	312.2	15.6	495	9	AI963720	AI963720 wf6604.x
18	311.4	15.6	3491	11	BC039100	BC039100 Homo sapi
C 19	311.2	15.6	925	9	AL558859	AL558859 AL558859
C 20	309.6	15.5	1910	11	BC035771	BC035771 Homo sapi
C 21	309.4	15.5	3491	11	BC039100	BC039100 Homo sapi
C 22	308.4	15.4	764	13	BU617735	BU617735 UI-H-DFO-
23	305.4	15.3	2076	11	BC035510	BC035510 Homo sapi
24	305	15.2	834	29	BZ612098	BZ612098 WHADP72TR
C 25	302.8	15.1	322	13	BU941317	BU941317 AGENCOURT
C 26	301.6	15.1	398	9	AL596113	AL596113 DKEZ6866A
C 27	301.2	15.1	868	10	BF337291	BF337291 602034658
C 28	299.8	15.0	1073	29	BZ601227	BZ601227 WHADP89TR
29	299.8	15.0	1910	11	BC035771	BC035771 Homo sapi
C 30	299.8	15.0	2076	11	BC035510	BC035510 Homo sapi
C 31	299.6	15.0	829	29	BZ609393	BZ609393 WHACI48TF
32	299.4	15.0	599	13	BU729004	BU729004 UI-E-CQ1-
C 33	299.2	15.0	691	9	AL043009	AL043009 DKEZp4340
C 34	298.8	14.9	584	12	BM701794	BM701794 UI-E-CQ1-
C 35	298.6	14.9	698	12	BM99040	BM99040 UI-H-DIO-
C 36	297.6	14.9	844	28	AQ748733	AQ748733 HS_5540.A
37	297.2	14.9	755	13	BX352666	BX352666 BX352666
38	297	14.8	957	9	AL573566	AL573566 AL573566
39	296	14.8	400	2	HS0072571	Bx482403 Homo sapi
C 40	296	14.8	910	13	BQ722917	BQ722917 AGENCOURT
C 41	295.8	14.8	540	13	BU617996	BU617996 UI-H-DFO-
C 42	295	14.8	564	10	BG529995	BG529995 602558836
43	294.8	14.7	815	29	BZ608327	BZ608327 WHACC28TR
C 44	294.6	14.7	533	28	AQ082371	AQ082371 RPL111-55
45	294.4	14.7	2284	11	BC036238	BC036238 Homo sapi

## ALIGNMENTS

RESULT 1	AF101960/c	2971 bp	DNA	linear	GSS 06-NOV-2000
LOCUS	AF101960	Human Homo sapiens	genomic clone pTWB59.14,	genomic survey	
DEFINITION	AF101960	Sequence.			
ACCESSION	AF101960	AF101960.1	GI:4193786		
VERSION	AF101960.1	GSS.			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 2971)				
AUTHORS	Bepler,G., O'Briant,K.C., Kim,Y.C., Schreiber,G. and Pittorle,D.M.				
TITLE	A 1.4-Mb high-resolution physical map and contig of chromosome segment 11p15.5 and genes in the LOH11A metastasis suppressor region				
JOURNAL	Genomics 55 (2), 164-175 (1999)				
MEDLINE	99134294				
PUBMED	9933563				
COMMENT	Contact: Bepler G				
	Medicine and Radiology				
	Duke University Medical Center				
	Box 2610, MSRB, Room 117, Durham, NC 27710, USA				
	part of a 1.4 megabase contig including the LOH11A metastasis suppressor region Bin T				
	Class: unknown.				



```

QY 939 AGTCACTTGAACCTGGAGCGAGGTTCAGTGGCGGAGATCATCATCACCCTCCCA 998
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1019 AATGGTGTGAACCCAGGAGCGAGCTTCAGTGGAGGAGATCATGCCACTGCATCCA 1078
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 999 GCGTGGGACAGACAGACTCTGCTCAATAAATAAATAAACAAGCAAGCAGT 1058
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1079 GCGTGGGACAGGCGAGACTCCGACTCAAAAAAAGATAGAAATGAACATG 1138
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1059 TTGTTGACTTATCTATAAAAAAATGCTCAACAAATAGAGCAGCAAGTGA 1118
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1139 GATATCTGCTACCTCATTTCTTCAACACTACTAGAAATACTCAAGGAATCA 1198
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1119 ATAAAGGAAATAAA 1133
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1199 GCAAGCAAAAAAGAA 1213
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 3
AL580585 929 bp mRNA linear EST 01-JUN-2003
LOCUS AL580585 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
DEFINITION Homo sapiens cDNA clone CS0DJ015YK20 3-PRIME, mRNA sequence.
ACCESSION AL580585
VERSION AL580585.2 GI:31318853
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 929)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
On Feb 16, 2001 this sequence version replaced gi:12946748.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paradise Avenue Genoscope sequence ID : CS0DJ015BF10NP1.
Location/Qualifiers
1. 929
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ015YK20"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 203 a 210 c 197 g 271 t 48 others
ORIGIN
Query Match 17.6%; Score 352.2; DB 9; Length 929;
Best Local Similarity 71.7%; Pred. No. 3.2e+02;
Matches 500; Conservative 6; Mismatches 174; Indels 17; Gaps 3;

QY 349 TTTTGTGTTTGTGAGACGAGTCTCGCTCTCGCCAGGCTGGA 408
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 105 TTTTGTGTTTGTGAGACGAGTCTCGCTCTCGCCAGGCTGGA 164
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 409 GTCACTGGCGGAGTCTCGCTCTCGCTCTCGCCAGGCTGGA 468
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 165 GCACAGTGGCGTGATVTCAGTCCGTGCAACCTCTGCTCCCGTTCAGCAATVTC 224
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 469 TGCCTCAGCTCCCAAGTAGCTGGGACTACAGGCGCGCCACTACGCCCGGCTAATTT 528
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 225 TGCCTCAGCTCCCGAGTGGCTGGGACTACAGGCGGTGCCACCACTCAGCTAA--TT 282
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 529 TTTGTATTTTACTAGAGAGCGGTTTACCGTTTACCGGTTTACCGGTTTACCGGTTT 588
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 283 TTTGTATTTTCCAGTAGAGAC--TGTTCACCATGTTGGTCAGCTGTGCTGAACCTG 340
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 589 ACCTCGTATCCCGCCCGCTCGCCCTCCCAAGTGTGGGATTTACAGGCGTGAGCCACG 648
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 341 ACTTCGTATCCCGCCCGCTCGCCCTCCCAAGTGTGGGATTTACAGGCGTGAGCCAC 400
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 649 CGCCCGGCGATGATCATCTCTTGTGACTATGCTGATGCAAGTACCTTAAGCCATC 708
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 401 CCCGGCGGAAGTTAAAGATGCACTTTGGGATAGACCATCATAGGATATTTCTCCCC 460
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 709 CTCTACCTTTTAAATATGCAAGTTTGGCCAGCCGCTGCTCATGCTGTAAATTCAGC 768
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 461 TTTTACTCATAAAGTTGTCAGCTTTGGCGGCGCGATGCTCAGCGCTGTATCTTAGT 520
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 769 ACTTTGGGAGGAGAGGTGGTGAATCACTTTGAGGCGAGAGTTTGAGACACCGCTGCC 828
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 521 ACTTTGGGAGGCTGAGGCGGCGAGATCACTGAGTCTGAGGAGTTTGTGACCGCTGCC 580
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 829 AACATGTAACCTCTGCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 888
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 581 AACATGTAACCTCTGCTTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 627
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 889 GTGCTGGGCGACACCTGTAATCCAGTATGCTGGAGGCTGAGGCGAGAGAGTCACTGA 948
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 628 GTGCTGGGCGACACCTGTAATCCAGTATGCTGGAGGCTGAGGCGAGAGAGTCACTGA 687
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 949 ACCTCGAGGCGAGGTTCAGTGGCGGAGATCATCATCAGCCCTCCAGCTGGCGGA 1008
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 688 ACCTGGGAGGAGAGGTTCAGTGGCGGAGATCATCATCAGCTCCAGCTGGCGGA 747
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 1009 CAGAGCAAGACTCTGCTCAATTAATAATAATAATAATAATAATAATAATAATA 1045
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 748 CAAGAGCAAGACTCTGCTCAATTAATAATAATAATAATAATAATAATAATAATA 784
  ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
AL120416/c
LOCUS AL120416 HEMBB1 Homo sapiens cDNA clone HEMBB1000637 5', mRNA
DEFINITION sequence.
ACCESSION AL120416
VERSION AL120416.1 GI:10935651
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 652)
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
HRI human cDNA project
Unpublished
Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
Location/Qualifiers

```

FEATURES













Db	2682	AGCCGAGATGGCCACCATTGCACCTCCAGCTGGGAGACAAGAGCAAACTCCGCTTACCA	2741
Qy	1032	AAATAAATAACAAACGAACAA	1053
Db	2742	AAAAAANAANAAGAAAAA	2763
RESULT 13			
CC061638/c			
LOCUS	CC061638	381 bp	DNA linear
DEFINITION	MUGO_CH352P00403T7_M05_CD318_034	CHORI-252 Vervet Monkey Library	GSS 28-APR-2003
		Cercopithecus aethiops genomic clone CH252-4B9, genomic survey	
		sequence.	
ACCESSION	CC061638		
VERSION			
KEYWORDS			
SOURCE	CC061638.1	GI:29790901	
ORGANISM	Cercopithecus aethiops (African green monkey)		
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
	1..381	Location/Qualifiers	
	/organism="Cercopithecus aethiops"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:9534"		
	/clone="CH252-4B9"		
	/sex="male"		
	/cell_type="white blood cell"		
	/dev_stage="Adult"		
	/clone_lib="CHORI-252 Vervet Monkey Library"		
	/notes="Vector: pTABAC3.1; Site_1: EORI; Site_2: EORI; Constructed by Michael Nefedov in Pieter de Jong's laboratory at BACPAC Resources, Children's Hospital in Oakland Research Institute."		
BASE COUNT	64 a	145 c	101 g
ORIGIN			71 t
Query Match	15.8%	Score 316.8	DB 29; Length 381;
Best Local Similarity	93.8%	Pred. No. 2.1e+03:	
Matches 330;	Conservative	0; Mismatches	22; Indels
		0; Gaps	0;
Qy	1649	GGAGACCCCTCCCATTCCTCGGGCCAGGAGGTAGGAGTGACATTCGCGACTGGGTGG	1708
Db	381	GGAGTCCCTCCCATCTCTGAGCCAGGAGGTAGGAGTGTATTCGCGACGGATGG	322
Qy	1709	GGGTGCTCTGCGGGTGAGATAGGGGACAGGAGGAGCTATTGCTAAGGCCGATAGG	1768
Db	321	GGGTGCTCTGCGGGTGAGATGGGGGACAGGAGGAGGATTCGTAAGGCCCAATAGG	262
Qy	1769	CACCTATTCCCGGGAATGCCCCAGGAGGAGTGGGTGTTATTAACTCAGSCCGGT	1828
Db	261	CACCTATGCCCCGGGAATGTCCCCAGGAGGAGTGGGTGTTATTAACTCAGSCCTGGT	202
Qy	1829	GCCCAGGAGCCAGGAGGAGGACGTGGGCCAGGAAGGACAGGCCCTGAGAAGTCTCGGCTG	1888

```

Db      201  GCCACGAGCCAGGAGGAGCGATGGCCACGAGAGTCCAGGCTTGAGAAATCCGCGGCTG 142
Qy      1889  AGCTGGGAGCAAAATCCCCCAACCCCTACCTTGGGGACAGGCAAGTGAGACCTGGTGAGG 1948
Db      141  AACTGGAAGCAATCCCCCAACCCCTACCTTGGGGACAGGCGAGGTGAGACCTGGGGAG 82
Qy      1949  GTGGCTCAGCAGGAGGAGGAGAGAGTGTCGTGGCTCTGCACCCACATC 2000
Db      81  GTGGCTCAGCAGGCGAGGAGAGAGGTGCTGTGGCTCTGCACCCACATC 30

RESULT 14
AG010128      709 bp      DNA      linear      GSS 14-APR-1999
LOCUS      Homo sapiens genomic DNA, 21q region, clone: f2G2x4, genomic survey
DEFINITION
ACCESSION
VERSION      AG010128.1 GI:3294404
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS      Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE      Homo sapiens genomic DNA, chromosome 21q
JOURNAL      Published Only in DataBase (1998)
REFERENCE
AUTHORS      Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL      Submitted (08-JUL-1998) Masahira Hattori, Kitasato University,
Department of Science, JST Sequencing Laboratory, Kitasato 1-15-1,
Sagamihara 228, Japan (E-mail:hattori@hg.ims.u-tokyo.ac.jp,
Tel:0427-78-9732, Fax:0427-78-9561)
COMMENT      On Feb 5, 1999 this sequence version replaced gi:2760798.
AG003939      Location/Qualifiers
FEATURES
source      1..709
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="21"
            /map="21q"
            /clone="f2G2x4"
BASE COUNT      171 a 179 c 184 g 168 t
ORIGIN
Query Match      15.8%; Score 315.6; DB 29; Length 709;
Best Local Similarity 73.3%; Pred. No. 1.4e+03;
Matches 499; Conservative 0; Mismatches 139; Indels 43; Gaps 6;

Qy      375  TGACAGGAGTCTCGTCTCTGCGCCAGGCTGGAGTCAGTGGCGGGATCTCGGCTCACT 434
Db      7  TGAATTCGCCNCCCGGGAGGTCGCCAGGTGGAGTCAGTGGCGGGATCTCGGCTCACT 66
Qy      435  GCAAGCTCCGCTCCCGGGTTCACGCCATCTCTCGCTCAGCCTCCCAAGTACTGGGA 494
Db      67  GCAGCTCCGCTCCCGGGTTCAGCCATATATCTCGCTCAACCTCCCAAGTACTGGGA 126
Qy      495  CTACAGGCGCCGCACTACGCCGGTAATTTTTTTGTATTTTTTAGTAGACGGGGTT 554
Db      127  CTACAGGCGCCGCACTACGCCGGTAATTTTTTTGTATTTTTTAGTAGACGGGGTT 186
Qy      555  TCACCGTTTAGCGGGGATGCCGTCGATCTCGTCACTCGTGATCCGCCGCCCTCGGCT 614
Db      187  TCACCATTTAGCGGGATGGTCTCGATCTCTTGACCTCGTGATCCGCCGCCCTCGGCT 246
Qy      615  CCCAAAGTGTGGGATACAGGCGTGAGCCACCGCGCGGCCATGATCATCTCTTGAC 674
Db      247  CCCAAAGTGTGGGATACAGGCGTGAG-CACCGCGCGGCCCTGAATAGTTTTTTGTG 305
Qy      675  TATGCTGATGTGACAAGTACCTAAAGCCATCAGACTCTACCCCTTTAATATGCAAGTTGG 734

```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 4, 2003, 20:39:54 ; Search time 511.264 Seconds  
(without alignments)  
10934.739 Million cell updates/sec

Title: US-09-936-271B-13\_COPY\_9500\_11570

Perfect score: 2071

Sequence: 1 aaacagacccaaacttctct.....cccgaaataaactgagaag 2071

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*

1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*
25:	/SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2063	99.6	11570	21	Human KLK-L2 gene.
C 2	482	23.3	586	24	Breast cancer rela
C 3	482	23.3	586	24	Breast cancer rela
C 4	482	23.3	735	21	cDNA sequence of h
C 5	482	23.3	735	24	Human breast tumou
C 6	482	23.3	738	20	DNA encoding huma
7	482	23.3	1302	20	Human BS247 specif
8	482	23.3	1302	20	Human BS247 specif

9	482	23.3	1381	20	AAZ06259	Human secreted pro
10	482	23.3	1439	20	AAZ06260	Human secreted pro
11	482	23.3	1499	20	AX60578	Human keratinocyte
12	482	23.3	1516	20	AAZ06245	Human secreted pro
13	482	23.3	1570	21	AAZ65070	Membrane-bound pro
14	482	23.3	1570	22	AAZ21471	Human cDNA sequenc
15	482	23.3	1570	22	AAF44216	Human PRO1132 (UNQ
16	482	23.3	1570	24	ABK28605	Human DNA59767-148
17	482	23.3	1570	25	ACA03830	cDNA encoding huma
18	482	23.3	1570	25	ACA04251	Human cDNA encodin
19	482	23.3	1570	25	ABX89368	DNA encoding novel
20	482	23.3	1570	25	ABX80318	Human secreted or
21	482	23.3	1570	25	ABX80822	Human secreted/tra
22	482	23.3	1570	25	ABX81205	Human secreted or
23	482	23.3	1570	25	ABX90295	Human secreted/tra
24	482	23.3	1570	25	ABX77906	Human PRO polynuc
25	482	23.3	1570	25	ABX79502	Human secreted/tr
26	482	23.3	1570	25	ABX64141	cDNA encoding huma
27	482	23.3	1570	25	ABX17105	Human PRO polynuc
28	481	23.2	1539	20	AAZ23317	Human PDSP-1 DNA.
29	476	23.0	1476	20	AAK16295	Human kallikrein e
30	460	22.2	1504	24	ABK30231	Human G-protein-co
31	449	21.7	673	22	AAF44901	Human breast cance
32	259	12.5	259	20	AAK57987	Human BS247 specif
33	251	12.1	254	20	AAK57988	Human BS247 specif
34	243.2	11.7	130263	24	ABK83573	Human cDNA differe
C 35	226	10.9	757	22	AAI96793	Human neuroblastom
C 36	225.4	10.9	240825	22	AAF24497	Human PG-3 gene.
C 37	225.4	10.9	240825	24	ABQ81802	Human PG-3 gene SE
C 38	222.6	10.7	11084	20	AAK22302	Human IL-1ra BAC c
C 39	219.4	10.6	70000	24	AAK42934	Human phospholipas
40	219.2	10.6	977	22	AAK82079	Human immune/haema
41	219.2	10.6	3830	20	AAK02984	Human IL-1ra BAC c
42	219	10.6	503	23	ABV54065	Human prostate exp
43	218.6	10.6	25464	19	AAV57274	Human flavin-conta
44	218.4	10.5	1446	22	AAH99110	Human EST-derived
45	218	10.5	4824	24	ABK70289	Human lung cancer

#### ALIGNMENTS

RESULT 1  
AAA95905  
ID AAA95905 standard; DNA; 11570 BP.  
XX  
AC AAA95905;  
XX  
DT 02-FEB-2001 (first entry)  
XX  
DE Human KLK-L2 gene.  
XX  
KW Human; KLK-L1; KLK-L2; KLK-L3; KLK-L4; KLK-L5; KLK-L6;  
KW Kallikrein-like protein; serine protease; cytosolic; cancer;  
KW prostrate cancer; ds.  
XX  
OS Homo sapiens.  
XX  
PN WO200053776-A2.  
XX  
PD 14-SEP-2000.  
XX  
PF 09-MAR-2000; 2000WO-CA00258.  
XX  
PR 11-MAR-1999; 99US-0124260.  
PR 01-APR-1999; 99US-0127386.  
PR 21-JUL-1999; 99US-0144919.  
PA (MOUN ) MOUNT SINAI HOSPITAL.  
XX  
PI Yousef GM, Diamandis EP;  
XX  
DR WPI; 2000-587440/55.

DR P-PSDB; AAB21296.

XX New kallikrein-like (KLK-L) proteins for diagnosing and treating KLK-L  
PT protein mediated disorders, especially cancer.

XX Claim 1; Page 143-149; 184pp; English.

XX The present sequence is the coding sequence of the human KLK-L2 gene,  
CC which encodes a kallikrein-like protein. Kallikreins and kallikrein-like  
CC proteins are a subgroup of the serine protease enzyme family. They  
CC catalyze the selective cleavage of specific polypeptide precursors to  
CC release peptides with potent biological activity. Nucleic acids encoding  
CC kallikrein-like proteins KLK-L1, KLK-L2, KLK-L3, KLK-L4, KLK-L5 and  
CC KLK-L6 have been isolated. The proteins are useful in the treatment,  
CC monitoring and diagnosis of cancers, especially prostate cancer. They  
CC can also be used to identify a substance that can associate with or  
CC mediate the biological activity of the proteins. Antibodies can be used  
CC to treat conditions mediated by the kallikrein-like proteins.

XX Sequence 11570 BP; 3033 A; 2562 C; 3326 G; 2649 T; 0 other;

Query Match 99.6%; Score 2063; DB 21; Length 11570;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2066; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AAACAGACCAAAACCTTCCTGCCCGTGGACCTCATGTTCCTCCCAAGTGGAGACAGCAA 60

DB 9500 AAACAGACCAAAACCTTCCTGCCCGTGGACCTCATGTTCCTCCCAAGTGGAGACAGCAA 9559

QY 61 TAAAGAGATAGATAAATATAGTAATTAATAAAAAAAAAAATAGCCGGGTGGTGG 120

DB 9560 TAAAGAGATAGATAAATATAGTAATTAATAAAAAAAAAAATAGCCGGGTGGTGG 9619

QY 121 CTTGCACCTGATGTTCCAGCTACCTGGAGGCTGAGGTGGAGAAATGCTTGACCCCAA 180

DB 9620 CTTGCACCTGATGTTCCAGCTACCTGGAGGCTGAGGTGGAGAAATGCTTGACCCCAA 9679

QY 181 CGTTTGAGGCTGCGGTAAAGCCATGACTGCTGCTGCTCCAGACAGCCCTGGGTGA 240

DB 9680 CGTTTGAGGCTGCGGTAAAGCCATGACTGCTGCTGCTCCAGACAGCCCTGGGTGA 9739

QY 241 CAAGCAGACGTTTTTGTGACAGAAAGAAAAAGAGAGAGAGAGAGAGAGAGAG 300

DB 9740 CAAGCAGACGTTTTTGTGACAGAAAGAAAAAGAGAGAGAGAGAGAGAGAGAG 9799

QY 301 AAAG 360

DB 9800 AAAG 9859

QY 361 GAAAG 420

DB 9860 GAAAG 9919

QY 421 GAGGAT 480

DB 9920 GAGGAT 9979

QY 481 TATTATAGGAGGTAATATAGGAGGATGAGGAGATGAGAGAGAGAGAGAGAGATAG 540

DB 9980 TATTATAGGAGGTAATATAGGAGGATGAGGAGATGAGAGAGAGAGAGAGATAG 10039

QY 541 TCCAGCGAATGGATTTCTATTGGAGGATGATTCGCCCTTGAAGACACTGCCAATACCA 600

DB 10040 TCCAGCGAATGGATTTCTATTGGAGGATGATTCGCCCTTGAAGACACTGCCAATACCA 10099

QY 601 GGAGACATTTTGGTGTACAACTATATGAGGGGCAATTAAGTGGCAATAGATAGA 660

DB 10100 GGAGACATTTTGGTGTGTACAACTATATGAGGGGCAATTAAGTGGCAATAGATAGA 10159

QY 661 TGCCAAAGTGCTGTTCAACATGCTATGATCAGACAGCGAGCGCTCCCAACAACCAATT 720

DB 10160 TGCCAAAGTGCTGTTCAACATGCTATGATCAGACAGCGAGCGCTCCCAACAACCAATT 10219

QY 721 ATCCAGCTTCAGATGCCAGTCCAGATCCAGAGAACCCCTCATCCAGGGGTGAGAAC 780  
DB 10220 ATCCAGCTTCAGATGCCAGTCCAGATCCAGAGAACCCCTCATCCAGGGGTGAGAAC 10279

QY 781 CGTATTTTTCAGAGGAGGTATAGGATGGTGGTGGAGATGGGGAAGAGGTGT 840  
DB 10280 CGTATTTTTCAGAGGAGGTATAGGATGGTGGTGGAGATGGGGAAGAGGTGT 10339

QY 841 GTGTCCAGTAAAGAGAAATAAGGCTGCACAGGCTGGAGGGGAGAGTCCAGAGAGAAAGGGA 900  
DB 10340 GTGTCCAGTAAAGAGAAATAAGGCTGCACAGGCTGGAGGGGAGAGTCCAGAGAGAAAGGGA 10399

QY 901 GCGGAGAGATACACGATGAGGGAGACAGGCTGGACAGAAAGTAGAGACAGATTCGA 960  
DB 10400 GCGGAGAGATACACGATGAGGGAGACAGGCTGGACAGAAAGTAGAGACAGATTCGA 10459

QY 961 GATGTGAGAGGAGGCTCAGACACCCCGGAAATGATGTGGACAAAGAGATTCGG 1020  
DB 10460 GATGTGAGAGGAGGCTCAGACACCCCGGAAATGATGTGGACAAAGAGATTCGG 10519

QY 1021 AAGAGGAAGATGAGTGGAGAGTGACAAATGGGCTCTAAAGGTTGAACCTTGGAGGCCAGG 1080  
DB 10520 AAGAGGAAGATGAGTGGAGAGTGACAAATGGGCTCTAAAGGTTGAACCTTGGAGGCCAGG 10579

QY 1081 CATGGTGGCTCAGCCCTGTATATCCCAACACTTTGGAGGCTGAGGTGGGCGAATCAGTTGA 1140  
DB 10580 CATGGTGGCTCAGCCCTGTATATCCCAACACTTTGGAGGCTGAGGTGGGCGAATCAGTTGA 10639

QY 1141 GCGCAGAGATTCGAGACAGCCTGGCCAAACATGTTGAAACCCCTCTCTACAAAAAAT 1200  
DB 10640 GCGCAGAGATTCGAGACAGCCTGGCCAAACATGTTGAAACCCCTCTCTACAAAAAAT 10599

QY 1201 ACAAAAAATAGCCGGGTGGTGGATGGACACCTGTAGTCACAGCTACTTTGGAGGCTGA 1260  
DB 10700 ACAAAAAATAGCCGGGTGGTGGATGGACACCTGTAGTCACAGCTACTTTGGAGGCTGA 10759

QY 1261 GCGAGAGAAATGCTTGAACCCGGAGATGGAGGCTGCAGTGAGCTGAGGTGAGGCCACT 1320  
DB 10760 GCGAGAGAAATGCTTGAACCCGGAGATGGAGGCTGCAGTGAGCTGAGGTGAGGCCACT 10819

QY 1321 GCGCTCAACCTGGGCAACAGAGTAAGACTCATCTCAAAAAAAGCTGATTTGG 1380  
DB 10820 GCGCTCAACCTGGGCAACAGAGTAAGACTCATCTCAAAAAAAGCTGATTTGG 10879

QY 1381 AGTGAATATTAATTAACATTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1440  
DB 10880 AGTGAATATTAATTAACATTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 10939

QY 1441 TTTTCTGCATTTCTTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500  
DB 10940 TTTTCTGCATTTCTTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 10999

QY 1501 TCCATCTATGGGCTCTCTGGTCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1560  
DB 11000 TCCATCTATGGGCTCTCTGGTCTCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 11059

QY 1561 GCCTCTCTCATGCCCT 1620  
DB 11060 GCCTCTCTCATGCCCT 11119

QY 1621 ATGGCTCCCTGAGGAGCTGCTGCT 1680  
DB 11120 ATGGCTCCCTGAGGAGCTGCTGCT 11179

QY 1681 CGGGTGTCTACACGAACCTCTGCAAGTTTCAAGTGGATCCAGAAACCAACCAACCAACCAACCAACCA 1740  
DB 11180 CGGGTGTCTACACGAACCTCTGCAAGTTTCAAGTGGATCCAGAAACCAACCAACCAACCAACCA 11239

QY 1741 ACTCCTGAGTATCCAGGAGCTCAGCACACCCGCAATCCCAACCTGCTGCTGAGGAGAGAGAGAGAG 1800  
DB 11240 ACTCCTGAGTATCCAGGAGCTCAGCACACCCGCAATCCCAACCTGCTGCTGAGGAGAGAGAGAG 11299

QY 1801 TGACACTCCTTTTACAGCCCTCATTTCTCCAGAGATGTTGAGAAATGTTTCTCTCTCTCTCTCTCTCT 1860



Qy 1950 CAAACTGTCCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 2009  
|||||  
Db 141 CAAACTGTCCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 82  
|||||  
Qy 2010 TCAGGGCCCATCCCTTCTCTGCGCTGTGACCAATTTAGTCCAGAAATAAAGTGAGA 2069  
|||||  
Db 81 TCAGGGCCCATCCCTTCTCTGCGCTGTGACCAATTTAGTCCAGAAATAAAGTGAGA 22  
|||||  
Qy 2070 AG 2071  
||  
Db 21 AG 20  
||  
RESULT 3  
ABL63996/C  
ID ABL63996 standard; DNA; 586 BP.  
XX AC ABL63996;  
XX DT 15-MAY-2002 (first entry)  
XX DE Breast cancer related gene sequence SEQ ID NO:2333.  
XX KW Human; cancer; colon; breast; ovary; oesophagus; kidney; thyroid;  
KW stomach; lung; prostate; pancreas; carcinoma; antitumour; cancerous;  
KW cytostatic; gene therapy; antineoplastic; Wilm's tumour; adenocarcinoma;  
KW gene; ds.  
XX OS Homo sapiens.  
XX PN WO200194629-A2.  
XX PD 13-DEC-2001.  
XX PF 30-MAY-2001; 2001WO-US10838.  
XX PR 05-JUN-2000; 2000US-209473P.  
PR 05-JUN-2000; 2000US-209531P.  
PR 18-SEP-2000; 2000US-233133P.  
PR 18-SEP-2000; 2000US-233617P.  
PR 20-SEP-2000; 2000US-234009P.  
PR 20-SEP-2000; 2000US-234034P.  
PR 20-SEP-2000; 2000US-234052P.  
PR 22-SEP-2000; 2000US-234509P.  
PR 22-SEP-2000; 2000US-234567P.  
PR 25-SEP-2000; 2000US-234923P.  
PR 25-SEP-2000; 2000US-234924P.  
PR 25-SEP-2000; 2000US-235077P.  
PR 25-SEP-2000; 2000US-235082P.  
PR 25-SEP-2000; 2000US-235134P.  
PR 26-SEP-2000; 2000US-235280P.  
PR 26-SEP-2000; 2000US-235637P.  
PR 26-SEP-2000; 2000US-235638P.  
PR 27-SEP-2000; 2000US-235711P.  
PR 27-SEP-2000; 2000US-235720P.  
PR 27-SEP-2000; 2000US-235840P.  
PR 27-SEP-2000; 2000US-235863P.  
PR 28-SEP-2000; 2000US-236028P.  
PR 28-SEP-2000; 2000US-236032P.  
PR 28-SEP-2000; 2000US-236033P.  
PR 28-SEP-2000; 2000US-236034P.  
PR 28-SEP-2000; 2000US-236109P.  
PR 29-SEP-2000; 2000US-236111P.  
PR 29-SEP-2000; 2000US-236842P.  
PR 29-SEP-2000; 2000US-236891P.  
PR 02-OCT-2000; 2000US-237172P.  
PR 02-OCT-2000; 2000US-237173P.  
PR 02-OCT-2000; 2000US-237278P.  
PR 02-OCT-2000; 2000US-237294P.  
PR 02-OCT-2000; 2000US-237295P.  
PR 02-OCT-2000; 2000US-237316P.  
PR 03-OCT-2000; 2000US-237423P.

03-OCT-2000; 2000US-237598P.  
PR 03-OCT-2000; 2000US-237604P.  
PR 03-OCT-2000; 2000US-237606P.  
PR 03-OCT-2000; 2000US-237608P.  
PR 01-NOV-2000; 2000US-244867P.  
PR 01-NOV-2000; 2000US-245084P.  
XX PA (AVAL-) AVALON PHARM.  
XX PI Young PE, Augustus M, Carter KC, Ebner R, Endress G, Horrigan S;  
PI Soppet DR, Weaver Z;  
XX DR WPI; 2002-188264/24.  
XX PT Screening for anti-neoplastic agent involves exposing cells to a  
PT chemical agent to be tested for anti-neoplastic activity, and  
PT determining a change in expression of a gene of a signature gene set  
XX  
XX PS Claim 1; SEQ ID 2333; 44pp; English.  
XX CC The present invention describes a method (M1) for screening for an  
CC anti-neoplastic agent. The method involves exposing cells to a chemical  
CC agent to be tested for anti-neoplastic activity, determining a change in  
CC expression of at least one gene (I) of a signature gene set, where (I)  
CC comprises a sequence (S) selected from 847 sequences (given in ABL61664  
CC to ABL70110), or is at least 95% identical to (S), where a change in  
CC expression is indicative of anti-neoplastic activity. (I) has cytostatic  
CC activity and can be used in gene therapy. M1 can be used for screening  
CC an anti-neoplastic agent, and can be used for producing a product which  
CC is the data collected with respect to the anti-neoplastic agent as a  
CC result of M1, and the data is sufficient to convey the chemical  
CC structure and/or properties of the agent. M1 can be used in the  
CC treatment of cancer such as colon, breast, stomach, lung, thyroid,  
CC oesophageal, ovarian, kidney, prostate or pancreatic cancer,  
CC adenocarcinoma, carcinoma, clear cell cancer, infiltrating ductal cancer,  
CC infiltrating lobular cancer, squamous cell carcinoma, neuroendocrine  
CC carcinoma, papillary carcinoma and Wilm's tumour.  
XX  
XX SQ Sequence 586 BP; 130 A; 139 C; 179 G; 136 T; 2 other;  
Query Match 23.3%; Score 482; DB 24; Length 586;  
Best Local Similarity 100.0%; Pred. No. 5.5e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1590 CAGGGTGATTCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGACATCGTGTCTCTGG 1649  
|||||  
Db 501 CAGGGTGATTCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGACATCGTGTCTCTGG 442  
|||||  
Qy 1650 GGAGATTACCTTGTGCGCGCGCCCAACAGACACGGGGTGTCTACAGAACCTCTGCAAGTTC 1709  
|||||  
Db 441 GGAGATTACCTTGTGCGCGCGCCCAACAGACACGGGGTGTCTACAGAACCTCTGCAAGTTC 382  
|||||  
Qy 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCAACTCCTGAGTATCCAGGACATCAGACACA 1769  
|||||  
Db 381 ACCAAGTGGATCCAGGAACCATCCAGGCCAACTCCTGAGTATCCAGGACATCAGACACA 322  
|||||  
Qy 1770 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCTCTTCAGACCCCTCATTCCTTC 1829  
|||||  
Db 321 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCTCTTCAGACCCCTCATTCCTTC 262  
|||||  
Qy 1830 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGACACTCAGG 1889  
|||||  
Db 261 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGACACTCAGG 202  
|||||  
Qy 1890 GTCTGCTTCCCGACATTTGGCTGACCGTGTCTCTCTAGTTGAACCTCGGAACAATTC 1949  
|||||  
Db 201 GTCTGCTTCCCGACATTTGGCTGACCGTGTCTCTCTAGTTGAACCTCGGAACAATTC 142  
|||||  
Qy 1950 CAAAACTGTCCAGGCGGGGTTGGGTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 2009  
|||||  
Db 141 CAAAACTGTCCAGGCGGGGTTGGGTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 82  
|||||  
Qy 2010 TCAGGGCCCATCCCTTCTCTGCGACTGTGACCCCAATTTAGTCCCAAAATAAAGTCTGAGA 2069

```
Db 81 TCAGGGCCCATCTCTCTGCGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 22
|||
QY 2070 AG 2071
||
Db 21 AG 20

RESULT 4
AAC79469/c
ID AAC79469 standard; cDNA; 735 BP.
XX AC
XX AAC79469;
XX 07-FEB-2001 (first entry)
XX cDNA sequence of human breast tumour clone B541S.
DE Human; breast tumour antigen; cytostatic; immunotherapy;
XX breast cancer; vaccine; ss.
XX Homo sapiens.
XX WO200061756-A2.
XX 19-OCT-2000.
XX 10-APR-2000; 2000WO-US09688.
XX 09-APR-1999; 99US-0288950.
XX 02-JUL-1999; 99US-0346327.
XX (CORI-) CORIXA CORP.
XX Reed SG, Xu J, Dillon DC;
PI WPI; 2000-638568/61.
XX
XX A novel isolated polypeptide comprising an immunogenic portion of a
PT breast cancer protein useful in the detection and treatment of breast
PT cancer -
XX
XX Claim 4; Page 88; 95pp; English.
XX
XX The present sequence was isolated from a breast tumour cDNA library. It
XX is provided in a specification relating to compounds for immunotherapy
XX and diagnosis of breast cancer. Breast tumour antigens and the
XX polynucleotides that encode them may be used in the production of a
XX pharmaceutical composition to be used in the treatment of breast cancer.
XX Proliferated T cells and incubated antigen presenting cells are also
XX required. The polypeptides and polynucleotides may also be used to
XX produce a vaccine.
XX
XX Sequence 735 BP; 161 A; 176 C; 226 G; 172 T; 0 other;
XX
Query Match 23.3%; Score 482; DB 21; Length 735;
Best Local Similarity 100.0%; Pred. No. 5.7e-81;
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1590 CAGGGTGATCTCGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTCTGG 1649
|||
Db 500 CAGGGTGATCTCGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTCTGG 441
|||
QY 1650 GGAGATTACCTTGTGCGGGGCCCAACAGACGCGGGTGTCTACAGAACCTCTGCAAGTTC 1709
|||
Db 440 GGAGATTACCTTGTGCGGGGCCCAACAGACGCGGGTGTCTACAGAACCTCTGCAAGTTC 381
|||
QY 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 1769
|||
Db 380 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTCTGAGTCATCCAGGACTCAGCACA 321
|||
QY 1770 CCGGCATCCCATCTGCTGAGGAGCAGCCCTCAGACTCTCTTTCAGACCTCATCTTC 1829
|||
```

```
Db 320 CCGGCATCCCATCTGCTGCGAGGACAGCCCTGACACTCTTTTCAGACCCCTCATCTTC 261
QY 1830 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGACTCAGG 1889
|||
Db 260 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGACTCAGG 201
|||
QY 1890 GTCTGCTTCCCCACATTTGGCTGACGCTGCTCTCTAGTTGAACCTGGGAACAATTC 1949
|||
Db 200 GTCTGCTTCCCCACATTTGGCTGACGCTGCTCTCTAGTTGAACCTGGGAACAATTC 141
|||
QY 1950 CAAAACTGTCAGGGGGGGTTCGCTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 2009
|||
Db 140 CAAAACTGTCAGGGGGGGTTCGCTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 81
|||
QY 2010 TCAGGGCCCATCTCTCTGCGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 2069
|||
Db 80 TCAGGGCCCATCTCTCTGCGAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTGA 21
|||
QY 2070 AG 2071
||
Db 20 AG 19

RESULT 5
ABK29013/c
ID ABK29013 standard; cDNA; 735 BP.
XX AC
XX ABK29013;
XX 23-APR-2002 (first entry)
XX Human breast tumour polypeptide cDNA clone #42.
DE Human; breast tumour polypeptide; gene; ss; breast cancer; cytostatic;
XX immunostimulant.
XX Homo sapiens.
XX WO200198339-A2.
XX 27-DEC-2001.
XX 12-JUN-2001; 2001WO-US19032.
XX 22-JUN-2000; 2000US-0602877.
XX 12-OCT-2000; 2000US-0687507.
XX 06-FEB-2001; 2001US-0778381.
XX (CORI-) CORIXA CORP.
XX Reed SG, Xu J, Dillon DC, Retter MW, Harlocker SL;
PI WPI; 2002-147792/19.
XX
XX Polynucleotides encoding breast tumour polypeptides, useful for treating
PT breast cancer or stimulating an immune response -
XX
XX Claim 1; Page 140; 150pp; English.
XX
XX The invention relates to polynucleotides encoding breast tumour
XX polypeptides. The sequences are useful for treating cancer, preferably
XX breast cancer, in a patient or for stimulating an immune response. The
XX polynucleotides and polypeptides are also useful in the diagnosis and
XX monitoring of breast cancer. A method for detecting the presence of a
XX cancer in a patient, comprises obtaining a biological sample from the
XX patient, contacting the biological sample with a binding agent that binds
XX to a breast tumour polypeptide, detecting in the sample an amount of
XX polypeptide that binds to the binding agent, and comparing the amount of
XX polypeptide to a predetermined cut-off value, therefore determining the
XX presence of a cancer in the patient. Sequences ABK28920-ABK29025
XX represent cDNA clones encoding human breast tumour polypeptides of the
XX invention.
```

Seq	Sequence	735 BP;	161 A;	176 C;	226 G;	172 T;	0 other;
	Query Match	23.3%;	Score 482;	DB 24;	Length 735;		
	Best Local Similarity	100.0%;	Pred. No. 5.7e-81;				
	Matches 482;	Conservative	0;	Mismatches	0;	Indels	0;
	Gaps	0;					
QY	1590	CAGGGTGATTC	GGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGACTCGTGTCTCTGG	1649			
Db	500	CAGGGTGATTC	GGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGACTCGTGTCTCTGG	441			
QY	1650	GGAGATTACCT	TGTGCCCGGGCCCAACAGACCCGGGTGTTACACGAACCTCTCTGCAAGTTC	1709			
Db	440	GGAGATTACCT	TGTGCCCGGGCCCAACAGACCCGGGTGTTACACGAACCTCTCTGCAAGTTC	381			
QY	1710	ACCAAGTGGATC	CAGGAACCAATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCAC	1769			
Db	380	ACCAAGTGGATC	CAGGAACCAATCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCAC	321			
QY	1770	CSGGCATCCCA	CTGCTGCAGGAGACGCTGACACTCCTTTTCAGACCCTCATTTCTCTTC	1829			
Db	320	CSGGCATCCCA	CTGCTGCAGGAGACGCTGACACTCCTTTTCAGACCCTCATTTCTCTTC	261			
QY	1830	CCAGAGATGTT	GAGAAATGTTTCATCTCTCCAGCCCCGACCCCATGTCTCTCGGACTCAGG	1889			
Db	260	CCAGAGATGTT	GAGAAATGTTTCATCTCTCCAGCCCCGACCCCATGTCTCTCGGACTCAGG	201			
QY	1890	GTCTGCTTCCC	CACATATGGGCTGACCGTGTCTCTTAGTTGAACCCCTGGGACAAATTC	1949			
Db	200	GTCTGCTTCCC	CACATATGGGCTGACCGTGTCTCTTAGTTGAACCCCTGGGACAAATTC	141			
QY	1950	CAAAACTGTCC	AGGGGGGGTTCGCTCTCAATCTCCCTGGGACATTTTCATCTCTCAAGC	2009			
Db	140	CAAAACTGTCC	AGGGGGGGTTCGCTCTCAATCTCCCTGGGACATTTTCATCTCTCAAGC	81			
QY	2010	TCAGGGCCCAT	CCCTTCTCTGCACTCTGACCCAAAATTTAGTCCGAAAAATAAACTGAGA	2069			
Db	80	TCAGGGCCCAT	CCCTTCTCTGCACTCTGACCCAAAATTTAGTCCGAAAAATAAACTGAGA	21			
QY	2070	AG	2071				
Db	20	AG	19				

RESULT 6	
AAx84240/c	
ID	AAx84240 standard; cDNA; 738 BP.
XX	
XX	
XX	AAx84240;
XX	
XX	08-SEP-1999 (first entry)
XX	
XX	DNA encoding human breast tumour protein immunogenic fragment.
DE	
XX	
XX	Breast tumour protein; immunogenic fragment; vaccine; detection;
KW	Breast cancer development; therapy; ss.
KW	
XX	
OS	Homo sapiens.
XX	
XX	W09933869-A2.
PN	
XX	
XX	08-JUL-1999.
PD	
XX	
XX	22-DEC-1998; 98WO-US27416.
PF	
XX	
XX	17-JUL-1998; 98US-0118627.
PR	
PR	24-DEC-1997; 97US-0998253.
PR	24-DEC-1997; 97US-0998255.
PR	17-JUL-1998; 98US-0118554.
XX	
XX	(CORI-) CORIXA CORP.
PA	
XX	
XX	Reed SG, Xu J;
PI	
XX	

DR WPI; 1999-405486/34.

XX New breast tumour protein genes used, in vaccines for immunotherapy,  
PT or for diagnosis of breast cancer

XX

XX PS PS PS

XX Claim 3; Page 70; 70pp; English.

XX

CC This sequence encodes a human breast tumour protein immunogenic fragment  
CC of the invention. The polypeptides or nucleic acids encoding them are  
CC useful in vaccines and pharmaceutical compositions for manufacture of  
CC medicaments for inhibiting the development of breast cancer in a patient.  
CC They can also be used to treat breast cancer. Antibodies against these  
CC polypeptides can be used to detect and monitor progression of breast  
CC cancer in patients. Primers and probes derived from the polynucleotides  
CC encoding the breast proteins are useful for detection of breast cancer.  
CC Peripheral blood cells from a patient incubated in the presence of at  
CC least one polypeptide, such that T cells proliferate, are useful in  
CC manufacture of a medicament for treating breast cancer in a patient.  
CC Antigen presenting cells incubated in the presence of at least one  
CC polypeptide are also useful for treating breast cancer.

XX

SQ Sequence 738 BP; 161 A; 179 C; 226 G; 172 T; 0 other;

Query Match 23.3%; Score 482; DB 20; Length 738;  
Best Local Similarity 100.0%; Pred. No. 5.7e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGTGATTCTCGGGGGCCCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1649  
DB 503 CAGGGTGATTCTCGGGGGCCCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 444  
QY 1650 GGAGATTACCTTTGTGCCCGGCCCAACAGACCAGGGGTGTCTACAGAACCTCTCGCAAGTTC 1709  
DB 443 GGAGATTACCTTTGTGCCCGGCCCAACAGACCAGGGGTGTCTACAGAACCTCTCGCAAGTTC 384  
QY 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCAACTCCTGAGTGCATCCAGGACTCAGCACA 1769  
DB 383 ACCAAGTGGATCCAGGAACCATCCAGGCCAACTCCTGAGTGCATCCAGGACTCAGCACA 324  
QY 1770 CCGGCATCCCCACCTGCTGCAGGACACAGCCCTGACACTCCTTTTCAGACCCTCATTTCTTC 1829  
DB 323 CCGGCATCCCCACCTGCTGCAGGACACAGCCCTGACACTCCTTTTCAGACCCTCATTTCTTC 264  
QY 1830 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGAGCTCAGG 1889  
DB 263 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGAGCTCAGG 204  
QY 1890 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTTTC 1949  
DB 203 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTCGGGAACAATTTTC 144  
QY 1950 CAAAACTGTCACAGGCGGGGTTGCGTCTCAATCTCCCTGGGACATTTTCATCCTCAAGC 2009  
DB 143 CAAAACTGTCACAGGCGGGGTTGCGTCTCAATCTCCCTGGGACATTTTCATCCTCAAGC 84  
QY 2010 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAAAATTTAGTCCCGAATAAATACTGAGA 2069  
DB 83 TCAGGGCCCATCCCTTCTCTGACGCTGACCCCAAAATTTAGTCCCGAATAAATACTGAGA 24  
QY 2070 AG 2071  
DB 23 AG 22

RESULT 7  
AA57989  
ID AA57989 standard; DNA; 1302 BP.  
XX  
AC AA57989;  
XX  
DT 19-JUL-1999 (first entry)  
XX  
DE Human BS247 specific polynucleotide #7.

RESULT 7	
AAK57980	
ID	AAK57989 standard; DNA; 1302 BP.
XX	
XX	
AC	AAK57989;
XX	
XX	
DT	19-JUL-1999 (first entry)
XX	
DE	Human BS247 specific polynucleotide #7.

XX BS247; detection; diagnosis; breast cancer; atypical hyperplasia;  
KW fibroadenoma; cystic breast disease; gene therapy; ss.  
XX Homo sapiens.  
XX WO9922027-A1.  
XX 06-MAY-1999.  
XX 28-OCT-1998; 98WO-US22906.  
XX 28-OCT-1997; 97US-0968838.  
XX 28-OCT-1997; 97US-0063431.  
XX (ABBO ) ABBOTT LAB.  
XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;  
PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;  
PI Russell JC, Stroupe SD, Yu H;  
XX WPI; 1999-312977/26.  
XX Breast tissue derived cDNA contig and consensus polypeptide sequence  
XX Claim 1; Page 105; 112pp; English.  
XX This sequence is a BS247 specific polynucleotide.  
XX The invention relates to a method of detecting the presence of a target  
XX BS247 polynucleotide, especially mRNA, in a test sample. BS247  
XX polynucleotides are derived from breast tissue. The polynucleotides,  
XX polypeptides or antibodies are useful for providing information leading  
XX to the detection, diagnosis, staging, monitoring, prognosis, in vivo  
XX imaging, prevention or treatment, determining predisposition to, diseases  
XX and conditions of the breast, such as breast cancer, atypical  
XX hyperplasia, fibroadenoma and cystic breast disease. Drug treatment or  
XX gene therapy for breast cancer, can be based on these identified gene  
XX sequences and the efficacy of any particular therapy can be monitored.  
XX The BS247-derived reagents are advantageous for detection of breast  
XX cancer due to their specificity. The reagents also provide an  
XX alternative, non-surgical diagnostic method capable of detecting early  
XX stage breast disease, such as cancer.  
XX Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 other;  
Query Match 23.3%; Score 482; DB 20; Length 1302;  
Best Local Similarity 100.0%; Pred. No. 6.1e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1590 CAGGGTGATTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG 1649  
DB 811 CAGGGTGATTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG 870  
QY 1650 GGAGATTACCTTGTGCCCCGGCCACACAGACCGGGGTGTACACGAACCTCTGCAAGTTC 1709  
DB 871 GGAGATTACCTTGTGCCCCGGCCACACAGACCGGGGTGTACACGAACCTCTGCAAGTTC 930  
QY 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTGTAGTCTATCCAGGACTCAGCACA 1769  
DB 931 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTGTAGTCTATCCAGGACTCAGCACA 990  
QY 1770 CCGGCATCCCACCTGTCTGAGGACAGCCCTGACACTCTCTTTCAGACCCCTCATCTCTTC 1829  
DB 991 CCGGCATCCCACCTGTCTGAGGACAGCCCTGACACTCTCTTTCAGACCCCTCATCTCTTC 1050  
QY 1830 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCTGGACTCAGG 1889  
DB 1051 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCTGGACTCAGG 1110  
QY 1890 GTCTGCTTCCCCACATTTGGGCTGACGGTGTCTCTAGTTGAACCCCTGGGAACAATTC 1949  
DB 1111 GTCTGCTTCCCCACATTTGGGCTGACGGTGTCTCTAGTTGAACCCCTGGGAACAATTC 1170

QY 1950 CAAAACGTGTCCAGGGGGGGTTCGCTCTCAATCTCCTCGGGCAGCTTTTCATCTCAAGC 2009  
DB 1171 CAAAACGTGTCCAGGGGGGGTTCGCTCTCAATCTCCTCGGGCAGCTTTTCATCTCAAGC 1230  
QY 2010 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTCTGAGA 2069  
DB 1231 TCAGGGCCCATCCCTTCTCTGCAGCTCTGACCCAAATTTAGTCCAGAAATAAAGTCTGAGA 1290  
QY 2070 AG 2071  
DB 1291 AG 1292  
RESULT 8  
AAX57990  
ID AAX57990 standard; DNA; 1302 BP.  
XX AAX57990;  
XX AC  
XX 19-JUL-1999 (first entry)  
XX DE Human BS247 specific polynucleotide #8.  
XX BS247; detection; diagnosis; breast cancer; atypical hyperplasia;  
KW fibroadenoma; cystic breast disease; gene therapy; ss.  
XX Homo sapiens.  
XX WO9922027-A1.  
XX 06-MAY-1999.  
XX 28-OCT-1998; 98WO-US22906.  
XX 28-OCT-1997; 97US-0968838.  
XX 28-OCT-1997; 97US-0063431.  
XX (ABBO ) ABBOTT LAB.  
XX Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;  
PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;  
PI Russell JC, Stroupe SD, Yu H;  
XX WPI; 1999-312977/26.  
XX Breast tissue derived cDNA contig and consensus polypeptide sequence  
XX Claim 1; Page 106; 112pp; English.  
XX This sequence is a BS247 specific polynucleotide.  
XX The invention relates to a method of detecting the presence of a target  
XX BS247 polynucleotide, especially mRNA, in a test sample. BS247  
XX polynucleotides are derived from breast tissue. The polynucleotides,  
XX polypeptides or antibodies are useful for providing information leading  
XX to the detection, diagnosis, staging, monitoring, prognosis, in vivo  
XX imaging, prevention or treatment, determining predisposition to, diseases  
XX and conditions of the breast, such as breast cancer, atypical  
XX hyperplasia, fibroadenoma and cystic breast disease. Drug treatment or  
XX gene therapy for breast cancer, can be based on these identified gene  
XX sequences and the efficacy of any particular therapy can be monitored.  
XX The BS247-derived reagents are advantageous for detection of breast  
XX cancer due to their specificity. The reagents also provide an  
XX alternative, non-surgical diagnostic method capable of detecting early  
XX stage breast disease, such as cancer.  
XX Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 other;  
Query Match 23.3%; Score 482; DB 20; Length 1302;  
Best Local Similarity 100.0%; Pred. No. 6.1e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1590 CAGGGTGATTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG 1649  
DB 811 CAGGGTGATTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCTGG 870  
QY 1650 GGAGATTACCTTGTGCCCCGGCCACACAGACCGGGGTGTACACGAACCTCTGCAAGTTC 1709  
DB 871 GGAGATTACCTTGTGCCCCGGCCACACAGACCGGGGTGTACACGAACCTCTGCAAGTTC 930  
QY 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTGTAGTCTATCCAGGACTCAGCACA 1769  
DB 931 ACCAAGTGGATCCAGGAACCATCCAGGCCAATCTGTAGTCTATCCAGGACTCAGCACA 990  
QY 1770 CCGGCATCCCACCTGTCTGAGGACAGCCCTGACACTCTCTTTCAGACCCCTCATCTCTTC 1829  
DB 991 CCGGCATCCCACCTGTCTGAGGACAGCCCTGACACTCTCTTTCAGACCCCTCATCTCTTC 1050  
QY 1830 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCTGGACTCAGG 1889  
DB 1051 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTCTGGACTCAGG 1110  
QY 1890 GTCTGCTTCCCCACATTTGGGCTGACGGTGTCTCTAGTTGAACCCCTGGGAACAATTC 1949  
DB 1111 GTCTGCTTCCCCACATTTGGGCTGACGGTGTCTCTAGTTGAACCCCTGGGAACAATTC 1170

Db 811 CAGGGTGAATTCCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTCTGG 870  
Qy 1650 GGAGATTACCTTGTGCGCGGCCCAACAGACACCGGGTGTCTTACAGAACCTCTGCAAGTTC 1709  
Db 871 GGAGATTACCTTGTGCGCGGCCCAACAGACACCGGGTGTCTTACAGAACCTCTGCAAGTTC 930  
Qy 1710 ACCAAGTGGATCCAGGAACCATCCAGGGCAACTCTGAGTCAATCCAGGACTCAGCACA 1769  
Db 931 ACCAAGTGGATCCAGGAACCATCCAGGGCAACTCTGAGTCAATCCAGGACTCAGCACA 990  
Qy 1770 CCGGCATCCCACTGCTGCAGGACAGCGCTGACACTCCTTTTCAGACCTCATTCCTTC 1829  
Db 991 CCGGCATCCCACTGCTGCAGGACAGCGCTGACACTCCTTTTCAGACCTCATTCCTTC 1050  
Qy 1830 CCAGAGATGTTGAGATGTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGAGCTCAGG 1889  
Db 1051 CCAGAGATGTTGAGATGTTCAATCTCTCCAGCCCTGACCCCATGTCTCTGAGCTCAGG 1110  
Qy 1890 GTCTGCTTCCCACTGAGCGGTGACCGTGTCTCTGAGTGAACCTGCGGAACAATTC 1949  
Db 1111 GTCTGCTTCCCACTGAGCGGTGACCGTGTCTCTGAGTGAACCTGCGGAACAATTC 1170  
Qy 1950 CAAACATGTCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACCTTCATCTCTCAAGC 2009  
Db 1171 CAAACATGTCAGGCGGGGTTGCGTCTCAATCTCCCTGGGCACCTTCATCTCTCAAGC 1230  
Qy 2010 TCAGGGCCCATCCCTTCTCTGAGCTGTGACCCAAATTTAGTCCCAAGAAATAAAGTGA 2069  
Db 1231 TCAGGGCCCATCCCTTCTCTGAGCTGTGACCCAAATTTAGTCCCAAGAAATAAAGTGA 1290  
Qy 2070 AG 2071  
Db 1291 AG 1292

RESULT 9  
AAZ06259  
ID AAZ06259 standard; DNA; 1381 BP.  
XX AC AAZ06259;  
XX AC AAZ06259;  
XX DT 30-SEP-1999 (first entry)  
XX DE Human secreted protein gene No. 27.  
XX KW Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
OS Homo sapiens.  
XX KW  
XX PN W09935158-A1.  
XX PD 15-JUL-1999.  
XX PF 06-JAN-1999; 99WO-US00108.  
XX PR 07-JAN-1998; 98US-0070704.  
XX PR 07-JAN-1998; 98US-0070657.  
XX PR 07-JAN-1998; 98US-0070658.  
XX PR 07-JAN-1998; 98US-0070692.  
XX PA (HUMA-) HUMAN GENOME SCI INC.  
XX PI Brewer LA, Duan RD, Ebner R, Lafleur DW, Ni J;  
PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;  
XX WPI; 1999-444190/37.

P-PSDB; AAY38426.  
New isolated human genes and the secreted polypeptides they encode  
Clalim 1; Page 173-174; 227pp; English.  
This sequence represents a nucleic acid molecule which encodes a secreted human protein. The gene number is given in the descriptor line. The gene can be used to generate fusion proteins by linking to the gene to a human immunoglobulin Fc portion (e.g. AAZ06210) for increasing the stability of the fused protein as compared to the human protein only. The invention relates to 36 novel genes and their fragments (nucleic acid sequences; AAZ06219-206263; amino acid sequences AAY38386-Y38498) which are useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. Also, pathological conditions can be diagnosed by determining the amount of the new polypeptides in a sample or by determining the presence of mutations in the new polynucleotides. Specific uses are described for each of the 36 polynucleotides, based on which tissues they are most highly expressed in (see AAZ06219 for described uses).  
XX Sequence 1381 BP; 329 A; 425 C; 353 G; 274 T; 0 other;  
XX SQ  
Query Match 23.3%; Score 482; DB 20; Length 1381;  
Best Local Similarity 100.0%; Pred. No. 6.1e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1590 CAGGGTGAATTCCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTCTGG 1649  
Db 852 CAGGGTGAATTCCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGCTCGTGTCTCTGG 911  
Qy 1650 GGAGATTACCTTGTGCGCGGCCCAACAGACACCGGGTGTCTTACAGAACCTCTGCAAGTTC 1709  
Db 912 GGAGATTACCTTGTGCGCGGCCCAACAGACACCGGGTGTCTTACAGAACCTCTGCAAGTTC 971  
Qy 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCAACTCTCTGAGTCAATCCAGGACTCAGCACA 1769  
Db 972 ACCAAGTGGATCCAGGAACCATCCAGGCCAACTCTCTGAGTCAATCCAGGACTCAGCACA 1031  
Qy 1770 CCGGCATCCCACTGCTGCAGGACAGCGCTGACACTCCTTTTCAGACCTCATTCCTTC 1829  
Db 1032 CCGGCATCCCACTGCTGCAGGACAGCGCTGACACTCCTTTTCAGACCTCATTCCTTC 1091  
Qy 1830 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGAGCTCAGG 1889  
Db 1092 CCAGAGATGTTGAGAAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGAGCTCAGG 1151  
Qy 1890 GTCTGCTTCCCACTGAGCGGTGACCGTGTCTCTTAGTTGAACCTGGGAACAATTC 1949  
Db 1152 GTCTGCTTCCCACTGAGCGGTGACCGTGTCTCTTAGTTGAACCTGGGAACAATTC 1211  
Qy 1950 CAAACATGTCAGGCGGGGGTTCGCTCAATCTCCCTGGGGGACTTTTCATCTCTCAAGC 2009  
Db 1212 CAAACATGTCAGGCGGGGGTTCGCTCAATCTCCCTGGGGGACTTTTCATCTCTCAAGC 1271  
Qy 2010 TCAGGGCCCATCCCTTCTCTGACCTCTGACCCAAATTTAGTCCCAAGAAATAAAGTGA 2069  
Db 1272 TCAGGGCCCATCCCTTCTCTGACCTCTGACCCAAATTTAGTCCCAAGAAATAAAGTGA 1331  
Qy 2070 AG 2071  
Db 1332 AG 1333  
RESULT 10  
AAZ06260  
ID AAZ06260 standard; DNA; 1439 BP.  
XX AC AAZ06260;  
XX AC AAZ06260;  
XX DT 30-SEP-1999 (first entry)  
XX DE Human secreted protein gene No. 27.



CC compositions. More generally the KDP polypeptide can be used in any  
CC cleaning composition, e.g. hard surface or dishwashing products.  
XX  
SQ Sequence 1499 BP; 310 A; 469 C; 398 G; 322 T; 0 other;

Query Match 23.3%; Score 482; DB 20; Length 1499;  
Best Local Similarity 100.0%; Pred. No. 6.2e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1590 CAGGGTGATTTCTGGGGGCGCTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1649  
Db 1014 CAGGGTGATTTCTGGGGGCGCTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1073  
Qy 1650 GGAGATTACCTTTGTGCGCGGCCCAACAGACCGGGGTCTTACAGAACCTCTGCAAGTTC 1709  
Db 1074 GGAGATTACCTTTGTGCGCGGCCCAACAGACCGGGGTCTTACAGAACCTCTGCAAGTTC 1133  
Qy 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCTGAGTCAATCCAGGACTCAGACACA 1769  
Db 1134 ACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCTGAGTCAATCCAGGACTCAGACACA 1193  
Qy 1770 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTTCAGACCTCATTCCTTC 1829  
Db 1194 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTTCAGACCTCATTCCTTC 1253  
Qy 1830 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1889  
Db 1254 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1313  
Qy 1890 GTCTGCTTCCCACTGCTGCAGGACAGCCCTGACACTCCTTTTCAGACCTCATTCCTTC 1949  
Db 1314 GTCTGCTTCCCACTGCTGCAGGACAGCCCTGACACTCCTTTTCAGACCTCATTCCTTC 1373  
Qy 1950 CAAAGTGTCCAGGCGGGGTGCGTCAATCTCCCTGCGGACATTCATCTCTCAAGC 2009  
Db 1374 CAAAGTGTCCAGGCGGGGTGCGTCAATCTCCCTGCGGACATTCATCTCTCAAGC 1433  
Qy 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCAGAGA 2069  
Db 1434 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCAGAGA 1493  
Qy 2070 AG 2071  
Db 1494 AG 1495

## RESULT 12

AZ06245  
ID AZ06245 standard; DNA; 1516 BP.  
XX  
AC AA206245;  
XX  
XX  
XX  
XX 30-SEP-1999 (first entry)  
XX  
DE Human secreted protein gene No. 27.  
XX  
XX Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
XX  
OS Homo sapiens.  
XX  
XX WO9935158-A1.  
XX  
XX 15-JUL-1999.  
XX  
XX 06-JAN-1999; 99WO-US00108.  
XX  
XX

PR 07-JAN-1998; 98US-0070704.  
PR 07-JAN-1998; 98US-0070657.  
PR 07-JAN-1998; 98US-0070658.  
PR 07-JAN-1998; 98US-0070692.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Brewer LA, Duan RD, Ebner R, Lafleur DW, NI J;  
PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;  
XX  
DR WPI; 1999-444190/37.  
DR P-PSDB; AAY38412.  
XX  
PT New isolated human genes and the secreted polypeptides they encode  
XX  
XX Claim 1; Page 163-164; 227pp; English.  
XX  
CC This sequence represents a nucleic acid molecule which encodes a  
CC secreted human protein. The gene number is given in the descriptor line.  
CC The gene can be used to generate fusion proteins by linking to the gene  
CC to a human immunoglobulin Fc portion (e.g. AA206210) for increasing the  
CC stability of the fused protein as compared to the human protein only.  
CC The invention relates to 36 novel genes and their fragments (nucleic  
CC acid sequences: AA206219-206263; amino acid sequences AAY38386-Y38498)  
CC which are useful for preventing, treating or ameliorating medical  
CC conditions e.g. by protein or gene therapy. Also, pathological  
CC polypeptides in a sample or by determining the amount of the new  
CC the new polynucleotides. Specific uses are described for each of the 36  
CC polynucleotides, based on which tissues they are most highly expressed in  
CC (see AA206219 for described uses).  
XX  
SQ Sequence 1516 BP; 338 A; 472 C; 381 G; 324 T; 1 other;

Query Match 23.3%; Score 482; DB 20; Length 1516;  
Best Local Similarity 100.0%; Pred. No. 6.2e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1590 CAGGGTGATTTCTGGGGGCGCTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1649  
Db 977 CAGGGTGATTTCTGGGGGCGCTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1036  
Qy 1650 GGAGATTACCTTTGTGCGCGGCCCAACAGACCGGGGTCTTACAGAACCTCTGCAAGTTC 1709  
Db 1037 GGAGATTACCTTTGTGCGCGGCCCAACAGACCGGGGTCTTACAGAACCTCTGCAAGTTC 1096  
Qy 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCTCTGAGTCAATCCAGGACTCAGACACA 1769  
Db 1097 ACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCTCTGAGTCAATCCAGGACTCAGACACA 1156  
Qy 1770 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCTCTCTAGTGAACCTCTGGGAACTTC 1829  
Db 1157 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCTCTCTAGTGAACCTCTGGGAACTTC 1216  
Qy 1830 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1889  
Db 1217 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1276  
Qy 1890 GTCTGCTTCCCACTGCTGCAGGACAGCCCTGACACTCTCTCTAGTGAACCTCTGGGAACTTC 1949  
Db 1277 GTCTGCTTCCCACTGCTGCAGGACAGCCCTGACACTCTCTCTAGTGAACCTCTGGGAACTTC 1336  
Qy 1950 CAAAGTGTCCAGGCGGGGTTCGCTCAATCTCCCTGCGGACATTTTCATCTCTCAAGC 2009  
Db 1337 CAAAGTGTCCAGGCGGGGTTCGCTCAATCTCCCTGCGGACATTTTCATCTCTCAAGC 1396  
Qy 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCAGAGA 2069  
Db 1397 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCAAATTTAGTCCAGAAATAAATCAGAGA 1456  
Qy 2070 AG 2071  
Db 1457 AG 1458

## RESULT 13

AAZ65070

ID AAZ65070 standard; cDNA; 1570 BP.

XX

AC AAZ65070:

XX

DT 05-APR-2000 (first entry)

XX

DE Membrane-bound protein Prol132 encoding cDNA.

XX

KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;

KW pharmaceutical; receptor immunoadhesin; gene mapping; ss.

XX

OS Homo sapiens.

XX

PN W09963088-A2.

XX

PD 09-DEC-1999.

XX

PF 02-JUN-1999; 99WO-US12252.

XX

PR 02-JUN-1998; 98US-0087607.

PR

PR 02-JUN-1998; 98US-0087609.

PR

PR 02-JUN-1998; 98US-0087759.

PR

PR 03-JUN-1998; 98US-0087827.

PR

PR 04-JUN-1998; 98US-0088021.

PR

PR 04-JUN-1998; 98US-0088025.

PR

PR 04-JUN-1998; 98US-0088028.

PR

PR 04-JUN-1998; 98US-0088029.

PR

PR 04-JUN-1998; 98US-0088030.

PR

PR 04-JUN-1998; 98US-0088033.

PR

PR 04-JUN-1998; 98US-0088326.

PR

PR 05-JUN-1998; 98US-0088167.

PR

PR 05-JUN-1998; 98US-0088202.

PR

PR 05-JUN-1998; 98US-0088212.

PR

PR 05-JUN-1998; 98US-0088217.

PR

PR 09-JUN-1998; 98US-0088655.

PR

PR 10-JUN-1998; 98US-0088722.

PR

PR 10-JUN-1998; 98US-0088730.

PR

PR 10-JUN-1998; 98US-0088734.

PR

PR 10-JUN-1998; 98US-0088738.

PR

PR 10-JUN-1998; 98US-0088740.

PR

PR 10-JUN-1998; 98US-0088741.

PR

PR 10-JUN-1998; 98US-0088742.

PR

PR 10-JUN-1998; 98US-0088810.

PR

PR 10-JUN-1998; 98US-0088811.

PR

PR 10-JUN-1998; 98US-0088824.

PR

PR 10-JUN-1998; 98US-0088825.

PR

PR 10-JUN-1998; 98US-0088826.

PR

PR 11-JUN-1998; 98US-0088858.

PR

PR 11-JUN-1998; 98US-0088861.

PR

PR 11-JUN-1998; 98US-0088863.

PR

PR 11-JUN-1998; 98US-0088876.

PR

PR 12-JUN-1998; 98US-0089090.

PR

PR 12-JUN-1998; 98US-0089105.

PR

PR 16-JUN-1998; 98US-0089440.

PR

PR 16-JUN-1998; 98US-0089512.

PR

PR 16-JUN-1998; 98US-0089514.

PR

PR 17-JUN-1998; 98US-0089532.

PR

PR 17-JUN-1998; 98US-0089538.

PR

PR 17-JUN-1998; 98US-0089598.

PR

PR 17-JUN-1998; 98US-0089600.

PR

PR 17-JUN-1998; 98US-0089653.

PR

PR 18-JUN-1998; 98US-0089801.

PR

PR 18-JUN-1998; 98US-0089907.

PR

PR 18-JUN-1998; 98US-0089908.

PR

PR 19-JUN-1998; 98US-0089947.

PR

PR 19-JUN-1998; 98US-0089948.

PR

PR 19-JUN-1998; 98US-0089952.

PR

PR 22-JUN-1998; 98US-0090246.

PR

PR 22-JUN-1998; 98US-0090252.  
PR 22-JUN-1998; 98US-0090254.  
PR 23-JUN-1998; 98US-0090349.  
PR 23-JUN-1998; 98US-0090355.  
PR 24-JUN-1998; 98US-0090429.  
PR 24-JUN-1998; 98US-0090431.  
PR 24-JUN-1998; 98US-0090435.  
PR 24-JUN-1998; 98US-0090444.  
PR 24-JUN-1998; 98US-0090445.  
PR 24-JUN-1998; 98US-0090461.  
PR 24-JUN-1998; 98US-0090472.  
PR 24-JUN-1998; 98US-0090535.  
PR 24-JUN-1998; 98US-0090538.  
PR 24-JUN-1998; 98US-0090540.  
PR 24-JUN-1998; 98US-0090557.  
PR 25-JUN-1998; 98US-0090676.  
PR 25-JUN-1998; 98US-0090678.  
PR 25-JUN-1998; 98US-0090688.  
PR 25-JUN-1998; 98US-0090690.  
PR 25-JUN-1998; 98US-0090691.  
PR 25-JUN-1998; 98US-0090694.  
PR 25-JUN-1998; 98US-0090695.  
PR 25-JUN-1998; 98US-0090696.  
PR 26-JUN-1998; 98US-0090862.  
PR 26-JUN-1998; 98US-0090863.  
PR 01-JUL-1998; 98US-0091358.  
PR 01-JUL-1998; 98US-0091360.  
PR 01-JUL-1998; 98US-0091344.  
PR 02-JUL-1998; 98US-0091478.  
PR 02-JUL-1998; 98US-0091486.  
PR 02-JUL-1998; 98US-0091519.  
PR 02-JUL-1998; 98US-0091626.  
PR 02-JUL-1998; 98US-0091628.  
PR 02-JUL-1998; 98US-0091633.  
PR 02-JUL-1998; 98US-0091646.  
PR 02-JUL-1998; 98US-0091673.  
PR 07-JUL-1998; 98US-0091978.  
PR 07-JUL-1998; 98US-0091982.  
PR 09-JUL-1998; 98US-0092182.  
PR 10-JUL-1998; 98US-0092472.  
PR 20-JUL-1998; 98US-0093339.  
PR 30-JUL-1998; 98US-0094651.  
PR 04-AUG-1998; 98US-0095282.  
PR 04-AUG-1998; 98US-0095285.  
PR 04-AUG-1998; 98US-0095301.  
PR 04-AUG-1998; 98US-0095302.  
PR 04-AUG-1998; 98US-0095318.  
PR 04-AUG-1998; 98US-0095321.  
PR 04-AUG-1998; 98US-0095325.  
PR 10-AUG-1998; 98US-0095916.  
PR 10-AUG-1998; 98US-0095929.  
PR 10-AUG-1998; 98US-0096012.  
PR 11-AUG-1998; 98US-0096143.  
PR 11-AUG-1998; 98US-0096146.  
PR 12-AUG-1998; 98US-0096329.  
PR 17-AUG-1998; 98US-0096757.  
PR 17-AUG-1998; 98US-0096766.  
PR 17-AUG-1998; 98US-0096768.  
PR 17-AUG-1998; 98US-0096773.  
PR 17-AUG-1998; 98US-0096791.  
PR 17-AUG-1998; 98US-0096867.  
PR 17-AUG-1998; 98US-0096891.  
PR 17-AUG-1998; 98US-0096894.  
PR 17-AUG-1998; 98US-0096895.  
PR 17-AUG-1998; 98US-0096897.  
PR 18-AUG-1998; 98US-0096949.  
PR 18-AUG-1998; 98US-0096950.  
PR 18-AUG-1998; 98US-0096959.  
PR 18-AUG-1998; 98US-0096960.  
PR 18-AUG-1998; 98US-0097022.  
PR 19-AUG-1998; 98US-0097141.  
PR 20-AUG-1998; 98US-0097218.  
PR 24-AUG-1998; 98US-0097661.

PR 26-AUG-1998; 98US-0097951.  
PR 26-AUG-1998; 98US-0097952.  
PR 26-AUG-1998; 98US-0097954.  
PR 26-AUG-1998; 98US-0097955.  
PR 26-AUG-1998; 98US-0097971.  
PR 26-AUG-1998; 98US-0097974.  
PR 26-AUG-1998; 98US-0097978.  
PR 26-AUG-1998; 98US-0097979.  
PR 26-AUG-1998; 98US-0097986.  
PR 26-AUG-1998; 98US-0098014.  
PR 31-AUG-1998; 98US-0098525.  
PR 16-SEP-1998; 98US-0100634.  
PR 12-JAN-1999; 99US-0115365.  
XX XX (GETH ) GENENTECH INC.  
XX XX Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;  
PI Wood WI, Yuan J;  
XX XX WPI; 2000-072883/06.  
DR P-PSDB; AAY66726.  
XX XX Membrane-bound proteins and related nucleotide sequences -  
XX XX Claim 2; Fig 225; 822pp; English.  
XX XX The invention provides membrane-bound PRO polypeptides and  
CC polynucleotides encoding them. The PRO sequences of the invention were  
CC identified based on extracellular domain homology screening. The PRO  
CC sequences have homology with proteins including LDL receptors, TIE  
CC ligands and various enzymes. The membrane-bound proteins and receptor  
CC molecules are useful as pharmaceutical and diagnostic agents. Receptor  
CC immunoadhesins, for instance, can be used as therapeutic agents to block  
CC receptor-ligand interactions. The membrane-bound proteins can also be  
CC employed for screening of potential peptide or small molecule inhibitors  
CC of the relevant receptor/ligand interaction. The PRO encoding sequences  
CC are useful as hybridization probes, in chromosome and gene mapping and in  
CC the generation of antisense RNA and DNA. PRO nucleic acid sequences  
CC will also be useful for the preparation of PRO polypeptides, especially  
CC by recombinant techniques.  
XX XX  
XX XX Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;  
SQ  
Query Match 23.3%; Score 482; DB 21; Length 1570;  
Best Local Similarity 100.0%; Pred. No. 6.2e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1590 CAGGGTGATCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTGG 1649  
DB 1077 CAGGGTGATCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTGG 1136  
QY 1650 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACACGAACCTCTGCAAGTTC 1709  
DB 1137 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTGTCTACACGAACCTCTGCAAGTTC 1196  
QY 1710 ACCAAGTGGATCCAGGAACACCATCCAGGCCAACTCTGAGTCATCCAGGACTCAGCACA 1769  
DB 1197 ACCAAGTGGATCCAGGAACACCATCCAGGCCAACTCTGAGTCATCCAGGACTCAGCACA 1256  
QY 1770 CCGCATCCCACTGTCTGAGGACAGCCCTGTACACTCTTTTCAGACCTCATCTCTTC 1829  
DB 1257 CCGCATCCCACTGTCTGAGGACAGCCCTGTACACTCTTTTCAGACCTCATCTCTTC 1316  
QY 1830 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1889  
DB 1317 CCAGAGATGTTGAGAATGTTTCATCTCTCCAGCCCTGACCCCATGTCTCTGGACTCAGG 1376  
QY 1890 GTCTGCTTCCCACTATGGGCTGACCGTGTCTCTAGTGTGAACCTCTGGGAACAATTC 1949  
DB 1377 GTCTGCTTCCCACTATGGGCTGACCGTGTCTCTAGTGTGAACCTCTGGGAACAATTC 1436  
QY 1950 CAAACATGTCAGGCGGGGTTGGTCTCAATCTCCCTGGGACATTCATCTCTCAAGC 2009  
|||||

Db 1437 CAAAACGTGTCCAGGGCGGGGTTGCTCTCAATCTCCCTGGGGCACTTTTCATCTCAAGC 1496  
QY 2010 TCAGGCCCCATCCCTTCTCTGACAGCTCTGACCCAAATTTAGTCCCAAGATAAATGAGA 2069  
|||||  
Db 1497 TCAGGCCCCATCCCTTCTCTGACAGCTCTGACCCAAATTTAGTCCCAAGATAAATGAGA. 1556  
QY 2070 AG 2071  
||  
Db 1557 AG 1558  
RESULT 14  
AAS21471  
ID AAS21471 standard; cDNA; 1570 BP.  
XX XX  
AC AAS21471;  
XX XX  
XX XX 24-OCT-2001 (first entry)  
XX XX Human cDNA sequence encoding for PRO1132 polypeptide.  
XX XX Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;  
KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.  
XX OS Homo sapiens.  
XX XX WO200140466-A2.  
XX XX 07-JUN-2001.  
XX XX 01-DEC-2000; 2000WO-US32678.  
PR 01-DEC-1999; 99WO-US28301.  
PR 01-DEC-1999; 99WO-US28634.  
PR 02-DEC-1999; 99WO-US28551.  
PR 02-DEC-1999; 99WO-US28584.  
PR 02-DEC-1999; 99WO-US28585.  
PR 09-DEC-1999; 99US-0170262.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 20-DEC-1999; 99WO-US30999.  
PR 30-DEC-1999; 99WO-US31243.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 20-MAR-2000; 2000WO-US07377.  
PR 21-MAR-2000; 2000WO-US07532.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 10-NOV-2000; 2000WO-US30873.  
XX XX (GETH ) GENENTECH INC.  
XX XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX WPI; 2001-408281/43.  
DR P-PSDB; AAU12399.  
XX XX  
PT Isolated, secretory and transmembrane PRO polypeptide used to detect  
other PRO polypeptides, link bioactive molecules to cells expressing

PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
PT lung, breast, prostate, cervical  
PS Claim 3; Fig 455; 813pp; English.  
XX  
CC AAS21244-AAS21518 encode for novel human secretory and transmembrane  
CC PRO polypeptides. The PRO polypeptides are useful to detect other  
CC PRO polypeptides, to link bioactive molecules to cells expressing  
CC PRO polypeptides, to modulate biological activities of cells expressing  
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
CC polypeptide expression in a cell sample to that in a control sample.  
CC Some of the 275 sequences are also useful to stimulate the release of  
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
CC proliferation or differentiation of chondrocytes, the proliferation or  
CC gene expression in pericyte cells, the release of proteoglycans from  
CC cartilage, the proliferation of inner ear utricular supporting cells or  
CC of T-lymphocytes, the release of a cytokine from peripheral blood  
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of  
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
CC molecules involved in binding interactions. The polynucleotides encoding  
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
CC transgenic or knock out animals and can be used in gene therapy.  
XX  
SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

Query Match 23.3%; Score 482; DB 22; Length 1570;  
Best Local Similarity 100.0%; Pred. No. 6.2e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1590 CAGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCTCAGGAGCTCTGTCCTGG 1649  
Db 1077 CAGGTGATTCGGGGGCTGTGGTCTGCAATGGCTCCTCAGGAGCTCTGTCCTGG 1136  
QY 1650 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1709  
Db 1137 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1196  
QY 1710 ACCAAGTGATCCAGGAACCATCCAGGCCAATCTCTAGTCTATCCAGGACTCAGCACA 1769  
Db 1197 ACCAAGTGATCCAGGAACCATCCAGGCCAATCTCTAGTCTATCCAGGACTCAGCACA 1256  
QY 1770 CCGGATCCCACTGCTCAGGGACAGCCCTGACACTCCTTTCAGACCTCATCTCCTTC 1829  
Db 1257 CCGGATCCCACTGCTCAGGGACAGCCCTGACACTCCTTTCAGACCTCATCTCCTTC 1316  
QY 1830 CCAGAGATGTTGAGAAATGTTATCTCTCCAGCCCTGACCCCATGTCTCTCTGGACTCAGG 1889  
Db 1317 CCAGAGATGTTGAGAAATGTTATCTCTCCAGCCCTGACCCCATGTCTCTCTGGACTCAGG 1376  
QY 1890 GTCTGCTTCCCAATGCGGTGACCGTGTCTCTAGTCTAGTCTGACCGTGGGAACATTTTC 1949  
Db 1377 GTCTGCTTCCCAATGCGGTGACCGTGTCTCTAGTCTAGTCTGACCGTGGGAACATTTTC 1436  
QY 1950 CAAAAGTCTCAGGGCGGGGGTGGTCTCAATCTCCCTGGGCACTTTCATCTCCTCAAGC 2009  
Db 1437 CAAAAGTCTCAGGGCGGGGGTGGTCTCAATCTCCCTGGGCACTTTCATCTCCTCAAGC 1496  
QY 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTCTGAGA 2069  
Db 1497 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAGTCTGAGA 1556  
QY 2070 AG 2071  
Db 1557 AG 1558

RESULT 15  
AAF44216  
ID AAF44216 standard; cDNA; 1570 BP.  
XX

AC AAF44216;  
XX  
DT 02-APR-2001 (first entry)  
XX  
DE Human PRO1132 (UN0570) nucleotide sequence SEQ ID NO:308.  
XX  
KW Human; secreted and transmembrane protein; PRO; cytostatic;  
KW cell death; cancer; chromosomal mapping; gene mapping; tissue typing;  
KW diagnostic assay; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO2000073454-A1.  
XX  
PD 07-DEC-2000.  
XX  
PF 30-MAR-2000; 2000WO-US08439.  
XX  
PR 02-JUN-1999; 99WO-US12252.  
PR 23-JUN-1999; 99US-0141037.  
PR 07-JUL-1999; 99US-0143048.  
PR 20-JUL-1999; 99US-0144758.  
PR 26-JUL-1999; 99US-0145698.  
PR 28-JUL-1999; 99US-0146222.  
PR 17-AUG-1999; 99US-0149396.  
PR 15-SEP-1999; 99WO-US21090.  
PR 08-OCT-1999; 99US-0158663.  
PR 30-NOV-1999; 99WO-US28313.  
PR 01-DEC-1999; 99WO-US28301.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 13-MAR-2000; 2000WO-US06884.  
PR 20-MAR-2000; 2000WO-US07377.  
XX  
(GETH ) GENENTECH INC.  
XX  
PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
PI Ferrara N, Fong S, Gerber H, Gerritsen ME, Goddard A, Godowski PJ;  
PI Grimaldi CJ, Gurney AL, Kijavini IJ, Napier MA, Pan J, Paoni NF;  
PI Roy MA, Stewart TA, Tamas D, Watanabe CK, Williams PM, Wood WI;  
PI Zhang Z;  
XX  
DR WPI: 2001-032160/04.  
DR P-PSDB; AAB65249.  
XX  
PT PRO polynucleotides used to produce polypeptides used to target  
PT bioactive molecules such as toxins, radiolabels or antibodies, to  
PT specific cells, to cause targeted cell death -  
XX  
PS Claim 2; Fig 225; 935pp; English.  
XX  
CC The present invention describes human secreted and transmembrane PRO  
CC proteins. The PRO proteins have cytostatic activity. The PRO proteins  
CC can be used for targeted delivery of bioactive molecules, such as  
CC toxins, radiolabels or antibodies, that cause cell death. PRO nucleotide  
CC sequences, and their fragments, can be used as hybridisation probes, in  
CC chromosomal and gene mapping, and in the generation of anti-sense RNA  
CC and DNA. They may also be used to produce transgenic animals which are  
CC used to develop and screen therapeutically useful reagents. The PRO  
CC nucleotide and protein sequence can be used for tissue typing and in  
CC treating cancer. Anti-PRO antibodies can be used in diagnostic assays.  
CC AAF44270 to AAF44470 represent PCR primers and hybridisation probes used  
CC in the isolation of human PRO sequences. AAF44087 to AAF44269 and  
CC AAB65154 to AAB65300 represent human PRO polynucleotide and protein

CC sequences given in the exemplification of the present invention.

XX  
SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;  
Query Match 23.3%; Score 482; DB 22; Length 1570;  
Best Local Similarity 100.0%; Pred. No. 6.2e-81;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1590 CAGGGTGATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1649  
Db 1077 CAGGGTGATCTGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGACTCGTGTCTCTGG 1136  
Qy 1650 GGAGATTACCTTGTGCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1709  
Db 1137 GGAGATTACCTTGTGCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1196  
Qy 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCAAGTCTCTGAGTCATCCAGGACTCAGCACA 1769  
Db 1197 ACCAAGTGGATCCAGGAACCATCCAGGCCAAGTCTCTGAGTCATCCAGGACTCAGCACA 1256  
Qy 1770 CCGGCATCCCACTCTCTGCAGGGACAGCCCTGACACTCTCTTCAGACCTCATTCCTTC 1829  
Db 1257 CCGGCATCCCACTCTCTGCAGGGACAGCCCTGACACTCTCTTCAGACCTCATTCCTTC 1316  
Qy 1830 CCAGAGATGTTGAGATGTTCTCTCCAGCCCTTGACCCCATGTCTCTGGACTCAGG 1889  
Db 1317 CCAGAGATGTTGAGATGTTCTCTCCAGCCCTTGACCCCATGTCTCTGGACTCAGG 1376  
Qy 1890 GTCTGCTTCCCAATGTTGGCTGACCGTGTCTCTAGTTGAACCTGGGAACAATTC 1949  
Db 1377 GTCTGCTTCCCAATGTTGGCTGACCGTGTCTCTAGTTGAACCTGGGAACAATTC 1436  
Qy 1950 CAAACTGTCCAGGGGGGTTGCGTCTCAATCTCCCTGGGCACTTTCATCTCAAGC 2009  
Db 1437 CAAACTGTCCAGGGGGGTTGCGTCTCAATCTCCCTGGGCACTTTCATCTCAAGC 1496  
Qy 2010 TCAGGGCCCATCCCTTCTCTGACAGCTCTGACCCAAATTTAGTCCAGAAATAAACTGAGA 2069  
Db 1497 TCAGGGCCCATCCCTTCTCTGACAGCTCTGACCCAAATTTAGTCCAGAAATAAACTGAGA 1556  
Qy 2070 AG 2071  
Db 1557 AG 1558

Search completed: August 5, 2003, 05:55:17  
Job time : 513.264 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 5, 2003, 05:12:55 ; Search time 134.302 Seconds  
(without alignments)  
6806.329 Million cell updates/sec

Title: US-09-936-271b-13\_COPY\_9500\_11570  
Perfect score: 2071  
Sequence: 1 aaacagacacaaacttctt.....cccagaaataaactgagaag 2071

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_NA.\*  
1: /cgn2.6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2.6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2.6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2.6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2.6/ptodata/1/ina/PctUS\_COMB.seq.\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	482	23.3	735	4	US-09-602-877A-94	Sequence 94, Appl
C 2	482	23.3	1570	4	US-09-996-243-308	Sequence 308, App
C 3	476	23.0	1476	2	US-08-824-874-2	Sequence 2, Appli
C 4	476	23.0	1476	3	US-09-210-084-2	Sequence 2, Appli
C 5	476	23.0	1476	4	US-09-764-762-2	Sequence 2, Appli
C 6	460	22.2	1504	4	US-09-280-116-1	Sequence 1, Appli
C 7	222.6	10.7	176373	3	US-09-128-153-17	Sequence 17, Appl
C 8	219.4	10.6	70000	4	US-09-851-896-3	Sequence 3, Appli
C 9	219.2	10.6	152331	3	US-09-128-155-16	Sequence 16, Appl
C 10	218.6	10.6	25464	4	US-09-326-480A-4	Sequence 4, Appli
C 11	216.4	10.4	461	4	US-09-404-879A-1	Sequence 1, Appli
C 12	216.4	10.4	461	4	US-09-404-879A-3	Sequence 3, Appli
C 13	216.4	10.4	461	4	US-09-338-933-1	Sequence 1, Appli
C 14	216.4	10.4	461	4	US-09-338-933-3	Sequence 3, Appli
C 15	216.4	10.4	461	4	US-09-215-681-1	Sequence 1, Appli
C 16	216.4	10.4	461	4	US-09-215-681-3	Sequence 3, Appli
C 17	215.6	10.4	3460	4	US-09-904-615-44	Sequence 44, Appl
C 18	214.6	10.4	1001	4	US-09-671-317-170	Sequence 170, App
C 19	214.6	10.4	1001	4	US-09-671-317-450	Sequence 450, App
C 20	214.6	10.4	41684	4	US-09-536-059-1	Sequence 1, Appli
C 21	212.6	10.3	4421	2	US-08-257-963B-9	Sequence 9, Appli
C 22	212.6	10.3	4421	4	US-08-367-841A-9	Sequence 9, Appli
C 23	212.6	10.3	4421	4	US-08-520-373D-6	Sequence 6, Appli
C 24	212.6	10.3	4421	5	PCT-US95-07201-9	Sequence 9, Appli
C 25	212.6	10.3	148567	4	US-09-801-876B-3	Sequence 3, Appli
C 26	212.2	10.2	99500	4	US-09-798-096-10	Sequence 10, Appl
C 27	210.4	10.2	38653	4	US-09-922-445-1	Sequence 1, Appli

C 28	209.8	10.1	55827	4	US-09-813-133A-3	Sequence 3, Appli
C 29	209.2	10.1	66804	4	US-09-740-041-3	Sequence 3, Appli
C 30	209	10.1	15328	2	US-08-888-497-33	Sequence 33, Appl
C 31	209	10.1	15328	4	US-09-362-230-33	Sequence 33, Appl
C 32	209	10.1	15328	5	PCT-US94-07926-33	Sequence 33, Appl
C 33	209	10.1	36651	4	US-09-738-894A-3	Sequence 3, Appli
C 34	209	10.1	36651	4	US-09-964-469-3	Sequence 3, Appli
C 35	209	10.1	49312	4	US-09-671-317-485	Sequence 485, App
C 36	207.4	10.0	6088	4	US-09-620-312D-150	Sequence 190, App
C 37	207.2	10.0	319608	4	US-09-539-333D-1	Sequence 1, Appli
C 38	207.2	10.0	319608	4	US-09-679-409-1	Sequence 1, Appli
C 39	207	10.0	9365	4	US-09-608-285A-8	Sequence 8, Appli
C 40	207	10.0	9365	4	US-09-350-836B-8	Sequence 8, Appli
C 41	207	10.0	9365	4	US-09-370-265-8	Sequence 8, Appli
C 42	207	10.0	9365	4	US-09-557-800C-8	Sequence 8, Appli
C 43	207	10.0	14581	4	US-08-520-373D-4	Sequence 4, Appli
C 44	207	10.0	14747	4	US-09-608-285A-42	Sequence 42, Appl
C 45	207	10.0	14747	4	US-09-557-800C-42	Sequence 42, Appl

ALIGNMENTS

RESULT 1  
US-09-602-877A-94/c  
; Sequence 94, Application US/09602877A  
; Patent No. 6432707  
; GENERAL INFORMATION:  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE OF INVENTION: AND DIAGNOSIS OF BREAST CANCER  
; FILE REFERENCE: 210121.446C5  
; CURRENT APPLICATION NUMBER: US/09/602,877A  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 94  
; LENGTH: 735  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-602-877A-94

Query Match	23.3%	Score 482;	DB 4;	Length 735;
Best Local Similarity	100.0%	Pred. No. 6.7e-92;		
Matches 482;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1590	CAGGGTGATTCCTGGGGGGCTGTGGTCTGCAATGGCTCGCTGAGGACTCGTGTCTGG 1649		
DB	500	CAGGGTGATTCCTGGGGGGCTGTGGTCTGCAATGGCTCGCTGAGGACTCGTGTCTGG 441		
QY	1650	GGAGATTACCTTGTGCGCGGCCCAACAGACGGGGTGTCTACACGAACCTCTGCAAGTTC 1709		
DB	440	GGAGATTACCTTGTGCGCGGCCCAACAGACGGGGTGTCTACACGAACCTCTGCAAGTTC 381		
QY	1710	ACCAAGTGATCCAGGAACCACTCAGGCAACTCTGTAGTCTATCCAGGACTCAGACA 1769		
DB	380	ACCAAGTGATCCAGGAACCACTCAGGCAACTCTGTAGTCTATCCAGGACTCAGACA 331		
QY	1770	CCGGATCCCCACCTGCTCGAGGAGAGCCCTGACACTCTTTCAGACCTCTATTCCTTC 1829		
DB	320	CCGGATCCCCACCTGCTCGAGGAGAGCCCTGACACTCTTTCAGACCTCTATTCCTTC 261		
QY	1830	CCAGAGATGTTGAGATGTTTCATCTCCAGCCCTGACCCCATGTCCTTGGACTCAGG 1889		
DB	260	CCAGAGATGTTGAGATGTTTCATCTCCAGCCCTGACCCCATGTCCTTGGACTCAGG 201		
QY	1890	GTCGTGTTCCCGACATTTGGGTGACCGGTGTCCTCTAGTTGAACCTGGGAACAATTC 1949		
DB	200	GTCGTGTTCCCGACATTTGGGTGACCGGTGTCCTCTAGTTGAACCTGGGAACAATTC 141		
QY	1950	CAAAATGTCTCAGGGGGGGTGTGGTCTCAATCTCCTCCCTGGGCACTTTCATCCTCAAGC 2009		

|||||  
140 CAAAACGTCTCAGGCGGGGTTCGCTCAATCTCCCTGGGCACTTTCATCTCAAGC 81  
2010 TCAGGCGCCATCCTTCTCTGACCTCTGACCAATTTAGTCCAGAAATAAATCTGAGA 2069  
|||||  
80 TCAGGCGCCATCCTTCTCTGACCTCTGACCAATTTAGTCCAGAAATAAATCTGAGA 21  
2070 AG 2071  
||  
20 AG 19

RESULT 2  
US-09-996-243-308  
; Sequence 308, Application US/09996243  
; Patent No. 6478825  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kijavlin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; TITLE OF INVENTION: Acids Encoding the Same  
; FILE REFERENCE: P2730P1C13  
; CURRENT APPLICATION NUMBER: US/09/996,243  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16  
; PRIOR APPLICATION NUMBER: 60/062250  
; PRIOR FILING DATE: 1997-10-17  
; PRIOR APPLICATION NUMBER: 60/065186  
; PRIOR FILING DATE: 1997-11-12  
; PRIOR APPLICATION NUMBER: 60/065311  
; PRIOR FILING DATE: 1997-11-13  
; PRIOR APPLICATION NUMBER: 60/066770  
; PRIOR FILING DATE: 1997-11-24  
; PRIOR APPLICATION NUMBER: 60/075945  
; PRIOR FILING DATE: 1998-02-25  
; PRIOR APPLICATION NUMBER: 60/078910  
; PRIOR FILING DATE: 1998-03-20  
; PRIOR APPLICATION NUMBER: 60/083322  
; PRIOR FILING DATE: 1998-04-28  
; PRIOR APPLICATION NUMBER: 60/084600  
; PRIOR FILING DATE: 1998-05-07  
; PRIOR APPLICATION NUMBER: 60/087106  
; PRIOR FILING DATE: 1998-05-28  
; PRIOR APPLICATION NUMBER: 60/087607  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087609  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087759  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/087827  
; PRIOR FILING DATE: 1998-06-02  
; PRIOR APPLICATION NUMBER: 60/088021  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088025  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088026  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088028  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088029  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088030  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088033  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088326  
; PRIOR FILING DATE: 1998-06-04  
; PRIOR APPLICATION NUMBER: 60/088167  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088202  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088212  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088217  
; PRIOR FILING DATE: 1998-06-05  
; PRIOR APPLICATION NUMBER: 60/088655  
; PRIOR FILING DATE: 1998-06-09  
; PRIOR APPLICATION NUMBER: 60/088734  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088738  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088742  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088810  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088824  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088826  
; PRIOR FILING DATE: 1998-06-10  
; PRIOR APPLICATION NUMBER: 60/088858  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088861  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/088876  
; PRIOR FILING DATE: 1998-06-11  
; PRIOR APPLICATION NUMBER: 60/089105  
; PRIOR FILING DATE: 1998-06-12  
; PRIOR APPLICATION NUMBER: 60/089440  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089512  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089514  
; PRIOR FILING DATE: 1998-06-16  
; PRIOR APPLICATION NUMBER: 60/089532  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089538  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089598  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089599  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089600  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089653  
; PRIOR FILING DATE: 1998-06-17  
; PRIOR APPLICATION NUMBER: 60/089801  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089907  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089908  
; PRIOR FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: 60/089947  
; PRIOR FILING DATE: 1998-06-19

;; PRIOR APPLICATION NUMBER: 60/089948  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/089952  
;; PRIOR FILING DATE: 1998-06-19  
;; PRIOR APPLICATION NUMBER: 60/090246  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090252  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090254  
;; PRIOR FILING DATE: 1998-06-22  
;; PRIOR APPLICATION NUMBER: 60/090349  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090355  
;; PRIOR FILING DATE: 1998-06-23  
;; PRIOR APPLICATION NUMBER: 60/090429  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090431  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090435  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090444  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090445  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090472  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090535  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090540  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090542  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090557  
;; PRIOR FILING DATE: 1998-06-24  
;; PRIOR APPLICATION NUMBER: 60/090676  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090678  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090690  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090694  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090695  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090696  
;; PRIOR FILING DATE: 1998-06-25  
;; PRIOR APPLICATION NUMBER: 60/090862  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/090863  
;; PRIOR FILING DATE: 1998-06-26  
;; PRIOR APPLICATION NUMBER: 60/091360  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091478  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091544  
;; PRIOR FILING DATE: 1998-07-01  
;; PRIOR APPLICATION NUMBER: 60/091519  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091626  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091633  
;; PRIOR FILING DATE: 1998-07-02  
;; PRIOR APPLICATION NUMBER: 60/091978  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/091982  
;; PRIOR FILING DATE: 1998-07-07  
;; PRIOR APPLICATION NUMBER: 60/092182  
;; PRIOR FILING DATE: 1998-07-09

Query Match 23.3%; Score 482; DB 4; Length 1570;  
Best Local Similarity 100.0%; Pred. No. 8e-92;  
Matches 482; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGTGAATTCGGGGGCGCTGTGGTCTGCAATGGCTCCCTCGAGGACTCGTGTCTCTGG 1649  
|||||  
DB 1077 CAGGGTGAATTCGGGGGCGCTGTGGTCTGCAATGGCTCCCTCGAGGACTCGTGTCTCTGG 1136  
|||||  
QY 1650 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTCTACACGAACTCTGCAAGTTTC 1709  
|||||  
DB 1137 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTCTACACGAACTCTGCAAGTTTC 1196  
|||||  
QY 1710 ACCAAGTGATCCAGGAAACCATCCAGGCAACTCTCTGAGTCAATCCAGGACTCAGCAC 1769  
|||||  
DB 1197 ACCAAGTGATCCAGGAAACCATCCAGGCAACTCTCTGAGTCAATCCAGGACTCAGCAC 1256  
|||||  
QY 1770 CCGCATCCCCACCTGCTCGAGGACAGCCCTGACACTCTCTTCAGACCCCTCATTCCTTC 1829  
|||||  
DB 1257 CCGCATCCCCACCTGCTCGAGGACAGCCCTGACACTCTCTTCAGACCCCTCATTCCTTC 1316  
|||||  
QY 1830 CCAGAGATGTTGAGAAATGTTTCATCTCCAGCCCTGACCCCATGTCTCTCTGACACTCAGG 1889  
|||||  
DB 1317 CCAGAGATGTTGAGAAATGTTTCATCTCCAGCCCTGACCCCATGTCTCTCTGACACTCAGG 1376  
|||||  
QY 1890 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTGGGAACAATTC 1949  
|||||  
DB 1377 GTCTGCTTCCCCACATTTGGGCTGACCGTGTCTCTCTAGTTGAACCTGGGAACAATTC 1436  
|||||  
QY 1950 CAAACTGTCCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 2009  
|||||  
DB 1437 CAAACTGTCCAGGCGGGGGTTCGCTCTCAATCTCCCTGGGCACTTTTCATCTCAAGC 1496  
|||||  
QY 2010 TCAGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCCGAGAAATAAAGTGA 2069  
|||||  
DB 1497 TCAGGCCCATCCCTTCTCTGACGCTCTGACCCCAATTTAGTCCCGAGAAATAAAGTGA 1556  
|||||  
QY 2070 AG 2071  
DB 1557 AG 1558  
|||

RESULT 3  
US-08-824-874-2  
; Sequence 2, Application US/08824874  
; Patent No. 5962300  
; GENERAL INFORMATION:  
; APPLICANT: Hillman, Jennifer L.  
; APPLICANT: Lal, Preeti  
; TITLE OF INVENTION: NOVEL KALLIKREIN  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/824,874  
; FILING DATE: Filed Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Billings, Lucy J.  
; REGISTRATION NUMBER: 36,749  
; REFERENCE/DOCKET NUMBER: PF-0252 US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-855-0555  
; TELEFAX: 415-845-4166  
; INFORMATION FOR SEQ ID NO: 2:

## SEQUENCE CHARACTERISTICS:

LENGTH: 1476 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: KERANOT02  
CLONE: 820694  
US-08-824-874-2

## Query Match

23.0%; Score 476; DB 2; Length 1476;

Best Local Similarity 100.0%; Pred. No. 1.4e-90;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGTGATTCCTGGGGGCGCTGGTCTGCAATGGCTCCCTGCAGGACTCGTGCTCTGG 1649  
DB 1001 CAGGGTGATTCCTGGGGGCGCTGGTCTGCAATGGCTCCCTGCAGGACTCGTGCTCTGG 1060  
QY 1650 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTTGCAGTTTC 1709  
DB 1061 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTTGCAGTTTC 1120  
QY 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCCTGAGTGTATCCAGGACTCAGCACA 1769  
DB 1121 ACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCCTGAGTGTATCCAGGACTCAGCACA 1180  
QY 1770 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1829  
DB 1181 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1240  
QY 1830 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1889  
DB 1241 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1300  
QY 1890 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTGTGAACCTGGGAACAATTC 1949  
DB 1301 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTGTGAACCTGGGAACAATTC 1360  
QY 1950 CAAACTGTCCAGGGCGGGGTTGCGGTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 2009  
DB 1361 CAAACTGTCCAGGGCGGGGTTGCGGTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 1420  
QY 2010 TCAGGGCCCATCCCTCTCTGACGCTGACCCCAAAATTTAGTCCAGAAATAACT 2065  
DB 1421 TCAGGGCCCATCCCTCTCTGACGCTGACCCCAAAATTTAGTCCAGAAATAACT 1476

## RESULT 4

US-09-210-084-2

Sequence 2, Application US/09210084

Patent No. 6197511

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Lal, Preeti

TITLE OF INVENTION: NOVEL KALLIKREIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/210,084

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/824, 874

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0252 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ. ID. NO. 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1476 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: KERANOT02

CLONE: 820694

US-09-210-084-2

Query Match 23.0%; Score 476; DB 3; Length 1476;

Best Local Similarity 100.0%; Pred. No. 1.4e-90;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGTGATTCCTGGGGGCGCTGGTCTGCAATGGCTCCCTGCAGGACTCGTGCTCTGG 1649  
DB 1001 CAGGGTGATTCCTGGGGGCGCTGGTCTGCAATGGCTCCCTGCAGGACTCGTGCTCTGG 1060  
QY 1650 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTTGCAGTTTC 1709  
DB 1061 GGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACCTCTTGCAGTTTC 1120  
QY 1710 ACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCCTGAGTGTATCCAGGACTCAGCACA 1769  
DB 1121 ACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCCTGAGTGTATCCAGGACTCAGCACA 1180  
QY 1770 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1829  
DB 1181 CCGGCATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTTCAGACCCCTCATTCCTTC 1240  
QY 1830 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1889  
DB 1241 CCAGAGATGTTGAGATGTTTCATCTCTCCAGCCCTGACCCCATGCTCTCTGGACTCAGG 1300  
QY 1890 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTGTGAACCTGGGAACAATTC 1949  
DB 1301 GTCTGCTTCCCGACATTTGGGCTGACCGTGTCTCTAGTGTGAACCTGGGAACAATTC 1360  
QY 1950 CAAACTGTCCAGGGCGGGGTTGCGGTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 2009  
DB 1361 CAAACTGTCCAGGGCGGGGTTGCGGTCTCAATCTCCCTGGGACATTTTCATCTCAAGC 1420  
QY 2010 TCAGGGCCCATCCCTCTCTGACGCTGACCCCAAAATTTAGTCCAGAAATAACT 2065  
DB 1421 TCAGGGCCCATCCCTCTCTGACGCTGACCCCAAAATTTAGTCCAGAAATAACT 1476

## RESULT 5

US-09-764-762-2

Sequence 2, Application US/09764762

Patent No. 6472195

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

APPLICANT: Lal, Preeti

TITLE OF INVENTION: NOVEL KALLIKREIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

```
/
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA: US/09/764,762
/ APPLICATION NUMBER: US/09/764,762
/ FILING DATE: 16-Jan-2001
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 09/210,084
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J.
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0252 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 415-855-0555
/ TELEFAX: 415-845-4166
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1476 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: KERANOT02
/ CLONE: 820694
/ SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-764-762-2

Query Match 23.0%; Score 476; DB 4; Length 1476;
Best Local Similarity 100.0%; Pred. No. 1.4e-90;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1590 CAGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTCAGGGACTCGTGCTCGG 1649
Db 1001 CAGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTCAGGGACTCGTGCTCGG 1060
QY 1650 GGAGATTACCTTGTGCGGGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1709
Db 1061 GGAGATTACCTTGTGCGGGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTTC 1120
QY 1710 ACCAAGTGGATCCAGGAACCACTCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCACA 1769
Db 1121 ACCAAGTGGATCCAGGAACCACTCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCACA 1180
QY 1770 CCGGATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTCAGACCCTCATTCCTTC 1829
Db 1181 CCGGATCCCACTGCTGCAGGACAGCCCTGACACTCCTTTCAGACCCTCATTCCTTC 1240
QY 1830 CCAGAGATCTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAGG 1889
Db 1241 CCAGAGATCTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAGG 1300
QY 1890 GTCTGCTCCCACTATGGGCTGACCGGTCTCTCTAGTTGAACCCCTGGGAACAATTC 1949
Db 1301 GTCTGCTCCCACTATGGGCTGACCGGTCTCTCTAGTTGAACCCCTGGGAACAATTC 1360
QY 1950 CAAAAGTCTCAGGGGGGGTGGTCTCAATCTCCCTGGGGCACTTTCATCCTCAAGC 2009
Db 1361 CAAAAGTCTCAGGGGGGGTGGTCTCAATCTCCCTGGGGCACTTTCATCCTCAAGC 1420
QY 2010 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAT 2065
Db 1421 TCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCAGAAATAAAT 1476

RESULT 6
US-09-280-116-1
; Sequence 1, Application US/09280116A
; Patent No. 6331427
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE OF INVENTION: AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/128,155
; EARLIER FILING DATE: 1998-08-03
; CURRENT APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0

/
/
/ APPLICANT: Robison, Keith E.
/ TITLE OF INVENTION: Nucleic Acid Molecules Encoding Human Protease Homologs
/ FILE REFERENCE: 5800-24, 035800/176965
/ CURRENT APPLICATION NUMBER: US/09/280,116A
/ CURRENT FILING DATE: 1999-03-26
/ NUMBER OF SEQ ID NOS: 268
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 1504
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: trypsin-like serine proteases
US-09-280-116-1

Query Match 22.2%; Score 460; DB 4; Length 1504;
Best Local Similarity 99.6%; Pred. No. 3.1e-87;
Matches 482; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 1590 CAGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTCAGGGACTCGTGCTCGT-G 1648
Db 1015 CAGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTCAGGGACTCGTGCTCGT 1074
QY 1649 GGGAGATTACCTTGTGCGGGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTT 1708
Db 1075 GGGAGATTACCTTGTGCGGGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAGTT 1134
QY 1709 CACCAAGTGGATCCAGGAACCACTCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCAC 1768
Db 1135 CACCAAGTGGATCCAGGAACCACTCCAGGCCAACTCCTGAGTCACTCCAGGACTCAGCAC 1194
QY 1769 ACCGGCATCCCACTGCTGCAGGGACAGCCCTGACACTCCTTTCAGACCCTCATTCCTT 1828
Db 1195 ACCGGCATCCCACTGCTGCAGGGACAGCCCTGACACTCCTTTCAGACCCTCATTCCTT 1254
QY 1829 CCCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAG 1888
Db 1255 CCCAGAGATGTTGAGAAATGTTCACTCTCCAGCCCTGACCCCATGCTCTCTGAGTCTCAG 1314
QY 1889 GGTCTGCTTCCCACTGCTGACCGTGTCTCTAGTTGAACCCCTGGGAACAATTT 1948
Db 1315 GGTCTGCTTCCCACTGCTGACCGTGTCTCTAGTTGAACCCCTGGGAACAATTT 1374
QY 1949 CCAAAAGTCTCCAGGGGGGGTGGTCTCAATCTCCCTGGGGCACTTTCATCCTCAAG 2008
Db 1375 CCAAAAGTCTCCAGGGGGGGTGGTCTCAATCTCCCTGGGGCACTTTCATCCTCAAG 1434
QY 2009 CTCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGT-CCCAGAAATAAATGA 2067
Db 1435 CTCAGGGCCCATCCCTTCTCTGAGCTCTGACCCCAATTTAGTCCCGCCAGAAATAAATGA 1494
QY 2068 GAAG 2071
Db 1495 GAAG 1498

RESULT 7
US-09-128-155-17
; Sequence 17, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; FILE OF INVENTION: AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/128,155
; EARLIER FILING DATE: 1998-08-03
; CURRENT APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
```

SEQ ID NO 17  
LENGTH: 176373  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(176373)  
OTHER INFORMATION: n = A,T,C or G  
US-09-128-155-17

Query Match 10.7%; Score 222.6; DB 3; Length 176373;  
Best Local Similarity 84.0%; Pred. No. 2.9e-37;  
Matches 263; Conservative 0; Mismatches 49; Indels 1; Gaps 1;

QY 1058 AAGAGTTGAACCTTGGAGGCGCAGCATGGTGGCTCAGCGCTGTATCCCAACACTTT-GGA 1116  
Db 127012 AAGAGAAAATTTATTTGGCCAGCGACGGCGCTCATCCCTATATCCACACTTTGGGA 127071

QY 1117 GGCTGAGGTGGCGCAATCACTTGAGGCCAGAGTTCGAGACCAGCCTGGCCAAACATGGTG 1176  
Db 127072 GGCTGAGGCGAGTGGATCACTCAGGTTCAGGAGTTCGAGACCAGCCTGGCCAAACATGGTG 127131

QY 1177 AAACCCCGTCTCTACAAAAAATACAAAAAATAGCCGGTGTGGTGGACACCTGT 1236  
Db 127132 AAACCCCATCTCTACTAAAAAATACAAAAAATAGCCAGTGTGGTGGACACCTGT 127191

QY 1237 AGTCACAGCTACTTGGAGGCTGAGGCGAGGAGAAATGCTTGAACCGGGAGATGGAGGCT 1296  
Db 127192 ANTCCAGCTACACGGGAGGCTGAGGCGAGAGATCGCTTGAACCGGGAGGAGGAT 127251

QY 1297 GCAGTGAGCTGAGTCAAGGCACTGCGCTCCACCTGGGCAACAGAGTAAGACTCCACT 1356  
Db 127252 GCAGTGAGCAAGTTCAAGCCACTGACCCAGCCTGGCCAGCAAGAGACTTTGTCT 127311

QY 1357 CAAAAAARAAA 1369  
Db 127312 CAAAAAARAAA 127324

RESULT 8  
US-09-851-896-3/C  
Sequence 3, Application US/09851896  
Patent No. 6410325  
GENERAL INFORMATION:  
APPLICANT: C. Frank Bennett  
APPLICANT: Susan M. Freier  
APPLICANT: Andrew T. Watt  
TITLE OF INVENTION: ANTISENSE MODULATION OF PHOSPHOLIPASE A2, GROUP VI (CA2+-INDEPENDENT)  
FILE REFERENCE: RTS-0220  
CURRENT FILING DATE: 2001-05-08  
NUMBER OF SEQ ID NOS: 89  
SEQ ID NO 3  
LENGTH: 70000  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
US-09-851-896-3

Query Match 10.6%; Score 219.4; DB 4; Length 70000;  
Best Local Similarity 85.9%; Pred. No. 1.1e-36;  
Matches 255; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

QY 1074 GGCAGGATGGTGGCTCAGCGCTGTATCCCAACACTTT-GGAGGCTGAGTGGGCGAA 1132  
Db 36372 GGCAGGATGGTGGCTCAGCGCTGTATCCCGCACTTTGGGAGGCGCAAGAGGCGAGA 36313

QY 1133 TCACCTGAGCGCAGGAGTTCGAGACCAAGCTGGCCAAACATGGTGAACCCCGCTCTTACA 1192  
Db 36312 TCACCTGAGTGGAGTTCGAGACCAAGCTGGCCAAACATGGTGAACCCCGCTCTACT 36253

QY 1193 AAAAAATACAAAAAATAGCCGGTGTGGTATGGACACTGTAGTCACAGCTACTTGG 1252

Db 36252 AAAAAACACAAAAAATAGCCGGGCTGGTGGCGGCGCTAGTCCAGCTACGCGG 36193  
QY 1253 GAGGCTGAGCGAGGAGAAATTTGAACCCGGGAGATGGAGGCTGCAGTGAGCTGAGGTC 1312  
Db 36192 GAGGCTGAGCGAGGAGAAATTTGAACCCGGGAGGAGCTTGCAGTGAGCGGAGATC 36133  
QY 1313 AGGCCACTCGCTCCCAACCTGGCAACAGACAGTAGACTCCATCTCAAAAAA 1369  
Db 36132 GCGCCACTGCATCCAGCCTGGCGGACAGACGAGACTCGCTCTCAAAAAA 36076

RESULT 9  
US-09-128-155-16  
Sequence 16, Application US/09128155  
Patent No. 6117654  
GENERAL INFORMATION:  
APPLICANT: Pan, Yang  
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY  
FILE REFERENCE: 09404/052001  
CURRENT APPLICATION NUMBER: US/09/128,155  
CURRENT FILING DATE: 1998-08-03  
EARLIER APPLICATION NUMBER: US 60/091,650  
EARLIER FILING DATE: 1998-07-02  
EARLIER APPLICATION NUMBER: US 60/054,646  
EARLIER FILING DATE: 1997-08-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 16  
LENGTH: 152331  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(152331)  
OTHER INFORMATION: n = A,T,C or G  
US-09-128-155-16

Query Match 10.6%; Score 219.2; DB 3; Length 152331;  
Best Local Similarity 81.4%; Pred. No. 1.4e-36;  
Matches 254; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1058 AAGAGTTGAACCTTGGAGGCGCAGCATGGTGGCTCACGCCCTGTATCCCAACACTTTGGAG 1117  
Db 58133 AAGAAAAAATTTATTTGGCCAGGCGCGCTCATCCCTATAATCCAGCACTTTGGGA 58192

QY 1118 GCTGAGGTGGCGAATCACTTTGAGGCGCAGAGTTCGAGACCGCCTGGCCAAACATGGTGA 1177  
Db 58193 GGCTGAGGCGAGTGCATCCCTGAGGTTCGAGAGTTCGAGACCGCCTGGCCAAACATGGTGA 58252

QY 1178 AACCCGCTCTCTACAAAAAATACAAAAAATAGCCGGTGTGGTGCATGGACACCTGTGA 1237  
Db 58253 AACCCATCTCTACTAAAAAATACAAAAAATAGCCAGGTGTGGTGGTGGGACCTGTGA 58312

QY 1238 GTCACAGCTACTTTGGGAGGCTGAGGCGAGAGAAATTTGTTGAACCCGGGAGATGGAGCTG 1297  
Db 58313 ATCCAGCTACACGGGAGGCTGAGGCGAGGAGAAATTCGTTCAACCCGGGAGGAGAAATG 58372

QY 1298 CAGTGAGCTGAGTTCAGGCGCCTGCGCTCCAACTGGGCAACAGAGTAAGACTCCATCTC 1357  
Db 58373 CAGTGAGCGCAAGTTCAAGCCACTGCACCCGAGCTGGGCAACAGAGCAAGACTTTGTCTC 58432

QY 1358 AAAAAAARAAA 1369  
Db 58433 CAAAAAARAAA 58444

RESULT 10  
US-09-326-480A-4  
Sequence 4, Application US/09326480A  
Patent No. 6551792  
GENERAL INFORMATION:

APPLICANT: Marta Blumenfeld  
APPLICANT: Ilija Tchoumakov  
APPLICANT: Henri-Jean Garchon  
APPLICANT: Jean-Francois Bach  
TITLE OF INVENTION: NUCLEOTIDE SEQUENCE WHICH ENCODES A  
TITLE OF INVENTION: FLAVIN MONOOXYGENASE, THE CORRESPONDING PROTEIN AND THEIR  
FILE REFERENCE: GENSSET.064C1  
CURRENT FILING DATE: 1999-06-04  
CURRENT FILING DATE: 1999-06-04  
PRIOR APPLICATION NUMBER: PCT/FR97/02226  
PRIOR FILING DATE: 1997-12-05  
PRIOR APPLICATION NUMBER: FR 96/15032  
PRIOR FILING DATE: 1996-12-06  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 25464  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE: UNSURE  
OTHER INFORMATION: genomic  
US-09-326-480A-4

Query Match 10.6%; Score 218.6; DB 4; Length 25464;  
Best Local Similarity 85.0%; Pred. No. 1.3e-36;  
Matches 256; Conservative 0; Mismatches 44; Indels 1; Gaps 1;  
QY 1070 TGGAGCCAGGATGTGTGCTCAGGCTGTATATCCCAACACTTTG-GAGGCTGAGGTGGG 1128  
Db 18484 TGCAGATTGGGATGTGTGCTCAGGCTGTATATCCCAACACTTTGTTGAGGCTGAGGCGGG 18543  
QY 1129 CGAATCACATTTAGGCCAGGATTCGAGACGACGCTGCGCAACATGTGTAAACCCCGCTCTC 1188  
Db 18544 CGGATTACCTGAGGTGAGGATTCGAGACGACGCTGCGCAACATGTGTAAACCCCGCTCTC 18603  
QY 1189 TACAAAAAATACAAAAAATAGCCGGTGTGTGATGGACACCTGTAGTCACAGCTAC 1248  
Db 18604 TACTAAAAATACAAAAAATAGTCAGTGTGTGCTGCACTGTATCCAGCTAC 18663  
QY 1249 TTGGGAGGCTGAGGAGGAGATTCCTTGAACCCGGGAGATGGAGCTGCAGTGAGCTGA 1308  
Db 18664 TCGGGAGGCTGAGGAGGAGATTCCTTGAATCAGGAGGAGTGGAGGCTGCAGTGAGCTGA 18723  
QY 1309 GGTACAGGCACCTGCGCTCCAACTGGGCAACAGAGTAAGACTCCATCTCAAAAAA 1368  
Db 18724 GATCGGCCACTGCACCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCTAGCT 18783  
QY 1369 A 1369  
Db 18784 A 18784

RESULT 11  
US-09-404-879A-1/c  
Sequence 1, Application US/09404879A  
Patent No. 6468546  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: King, Gordon E.  
APPLICANT: Algate, Paul A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.462C2  
CURRENT APPLICATION NUMBER: US/09/404,879A  
CURRENT FILING DATE: 1999-09-24  
NUMBER OF SEQ ID NOS: 393  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 1  
LENGTH: 461  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-404-879A-1

Query Match 10.4%; Score 216.4; DB 4; Length 461;  
Best Local Similarity 84.4%; Pred. No. 1.5e-36;  
Matches 255; Conservative 0; Mismatches 46; Indels 1; Gaps 1;  
QY 1069 TTGGAGCCAGGATGTGTGCTCAGGCTGTATATCCCAACACTTTT-GGAGGCTGAGGTGG 1127  
Db 368 TTGGGGCGGGCGGTGTGTGCTCAGGCTGTATATCCCAACACTTTTGGGAGGCGGCGG 309  
QY 1128 GCGAATCACTTTGAGGCCAGGAGTTTCAGACGACGCTGGCCCAACATGTTGAAACCCGCTCT 1187  
Db 308 GTGGATCACTTGTGAGGCTGAGGTTCAAGAGCAGCTGGCCCAACCTGTTGAAACCCGCTCT 249  
QY 1188 CTACAAAAAATACAAAAAATAGCCGGTGTGTGATGGACACCTGTAGTCACAGCTA 1247  
Db 248 CTACTAAAAATACAAAAAATAGCTGAGCGTGTGTGGCGGCGCTGTATATCCAGCTA 189  
QY 1248 CTTGGAGGCTGAGGAGGAGATTCCTTGAACCCGGGAGATGGAGGCTGCAGTGAGCTG 1307  
Db 188 CTTGGAGGCTGAGGAGGAGATTCCTTGAACCTGGGAGGCGGAGGTTGCAGCGAGCTG 129  
QY 1308 AGTCAAGGCTGAGGCTGAGGAGGAGATTCCTTGAACCTGGGAGGCGGAGGTTGCAGCGAGCTG 1367  
Db 128 AGATCATGCCGTGTGTACTCCAGCTTGGGCAACAGAGTGAGACTCCATCTCAAAAAA 69  
QY 1368 AA 1369  
Db 68 AA 67

RESULT 12  
US-09-404-879A-3/c  
Sequence 3, Application US/09404879A  
Patent No. 6468546  
GENERAL INFORMATION:  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: King, Gordon E.  
APPLICANT: Algate, Paul A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
TITLE OF INVENTION: DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.462C2  
CURRENT APPLICATION NUMBER: US/09/404,879A  
CURRENT FILING DATE: 1999-09-24  
NUMBER OF SEQ ID NOS: 393  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 3  
LENGTH: 461  
TYPE: DNA  
ORGANISM: Homo sapien  
US-09-404-879A-3

Query Match 10.4%; Score 216.4; DB 4; Length 461;  
Best Local Similarity 84.4%; Pred. No. 1.5e-36;  
Matches 255; Conservative 0; Mismatches 46; Indels 1; Gaps 1;  
QY 1069 TTGGAGCCAGGATGTGTGCTCAGGCTGTATATCCCAACACTTTT-GGAGGCTGAGGTGG 1127  
Db 368 TTGGGGCGGGCGGTGTGTGCTCAGGCTGTATATCCCAACACTTTTGGGAGGCGGCGG 309  
QY 1128 GCGAATCACTTTGAGGCCAGGAGTTTCAGACGACGCTGGCCCAACATGTTGAAACCCGCTCT 1187  
Db 308 GTGGATCACTTGTGAGGCTGAGGTTCAAGAGCAGCTGGCCCAACCTGTTGAAACCCGCTCT 249  
QY 1188 CTACAAAAAATACAAAAAATAGCCGGTGTGTGATGGACACCTGTAGTCACAGCTA 1247  
Db 248 CTACTAAAAATACAAAAAATAGCTGAGCGTGTGTGGCGGCGCTGTATATCCAGCTA 189  
QY 1248 CTTGGAGGCTGAGGAGGAGATTCCTTGAACCCGGGAGATGGAGGCTGCAGTGAGCTG 1307  
Db 188 CTTGGAGGCTGAGGAGGAGATTCCTTGAACCTGGGAGGCGGAGGTTGCAGCGAGCTG 129  
QY 1308 AGTCAAGGCTGAGGCTGAGGAGGAGATTCCTTGAACCTGGGAGGCGGAGGTTGCAGCGAGCTG 1367  
Db 128 AGATCATGCCGTGTGTACTCCAGCTTGGGCAACAGAGTGAGACTCCATCTCAAAAAA 69





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 5, 2003, 09:57:15 ; Search time 3128 Seconds  
(without alignments)  
3832.008 Million cell updates/sec

Title: US-09-936-271B-14

Perfect score: 1608

Sequence: 1 MATARPPWMLCALITALL.....VYTNLCKFTKWIQTQANS 293

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q/Cgn2\_1/USPTO.spool/US09936271/runat\_04082003\_104245\_25702/app\_query.fasta\_1.455  
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09936271 @CGN\_1\_1.2326 @runat\_04082003\_104245\_25702 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl :  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pln.\*  
35: em.htg\_rod.\*  
36: em.htg\_mam.\*  
37: em.htg\_vrt.\*  
38: em.sy.\*  
39: em.htgo\_hum.\*  
40: em.htgo\_mus.\*  
41: em.htgo\_other.\*

pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match %	Length	DB	ID	Description
1	1608	100.0	882	9	BT006867	BT006867 Homo sapi
2	1608	100.0	882	12	BT007831	BT007831 Synthetic
3	1608	100.0	1370	9	AY279380	AY279380 Homo sapi
4	1608	100.0	1381	6	BD107879	BD107879 36 human
5	1608	100.0	1387	9	AF168768	AF168768 Homo sapi
6	1608	100.0	1438	9	AY279381	AY279381 Homo sapi
7	1608	100.0	1499	6	BD005362	BD005362 Protease.
8	1608	100.0	1516	6	BD107865	BD107865 36 human
9	1608	100.0	1536	9	BC008036	BC008036 Homo sapi
10	1608	100.0	1570	6	AR252580	AR252580 Sequence
11	1608	100.0	1570	6	AX080829	AX080829 Sequence
12	1608	100.0	1570	6	AX03421	AX03421 Sequence
13	1608	100.0	1570	6	AX464322	AX464322 Sequence
14	1595	99.2	1476	6	AR078184	AR078184 Sequence
15	1595	99.2	1476	6	AR137506	AR137506 Sequence
16	1595	99.2	1476	6	AR242358	AR242358 Sequence
17	1595	99.2	1476	6	BD082659	BD082659 Keratinoc
18	1533	94.1	1439	6	BD107880	BD107880 36 human
19	1475	91.7	929	9	AF435981	AF435981 Homo sapi
20	1475	91.7	1132	9	AF435980	AF435980 Homo sapi
21	1403	87.3	1504	6	AR263823	AR263823 Sequence
22	1002	62.3	11570	9	AF135028	AF135028 Homo sapi
23	1002	62.3	217346	2	AC027602	AC027602 Homo sapi
24	1002	62.3	230000	9	AF243527	AF243527 Homo sapi
25	997	62.0	107487	9	AC011483	AC011483 Homo sapi
26	985.5	61.3	200792	2	AC130782	AC130782 Pan trogl
27	725	45.1	735	6	AX429955	AX429955 Sequence
28	725	45.1	738	6	BD139877	BD139877 Compounds
29	688	42.8	673	6	AX067353	AX067353 Sequence
30	676	42.0	1140	4	SSU76256	U76256 Sus scrofa
31	671	41.7	765	6	AR261054	AR261054 Sequence
32	671	41.7	765	6	AR278585	AR278585 Sequence
33	671	41.7	765	6	AX141034	AX141034 Sequence
34	671	41.7	765	6	AX200894	AX200894 Sequence
35	671	41.7	765	6	AX267550	AX267550 Sequence
36	671	41.7	765	9	AF259969	AF259969 Homo sapi
37	670	41.7	1347	9	AF113140	AF113140 Homo sapi
38	654.5	40.7	1052	6	AR219287	AR219287 Sequence
39	652.5	40.6	1166	6	AR152173	AR152173 Sequence
40	650.5	40.5	833	6	AR060847	AR060847 Sequence
41	650.5	40.5	833	6	BD082136	BD082136 Novel pro
42	650.5	40.5	1106	6	AX016287	AX016287 Sequence
43	650.5	40.5	1106	6	BD137019	BD137019 Human CAS
44	650.5	40.5	1158	6	AX016289	AX016289 Sequence
45	650.5	40.5	1158	6	BD137020	BD137020 Human CAS

# ALIGNMENTS

RESULT 1



## AUTHORS

Kalnine,N., Chen,X., Rolfs,A., Halleck,A., Hines,L., Eisenstein,S.,  
Koundinya,M., Raphael,J., Moreira,D., Kelley,T., LaBaer,J., Lin,Y.,  
Phelan,M. and Farmer,A.

TITLE  
JOURNAL

Direct Submission

Submitted (13-MAY-2003) BD Biosciences Clontech, 1020 East Meadow  
Circle, Palo Alto, CA 94303, USA

## COMMENT

This CDS clone is a part of a collection of human full length  
expression clones generated by BD Biosciences Clontech and the  
Harvard Institute of Proteomics. Each CDS has been cloned in two  
forms: with and without stop-codon (to allow fusion with C-terminal  
tag). The CDS has been directionally cloned using BD In-Fusion(TM)  
cloning system between the SalI and HindIII sites of the pDNR-DUAL  
vector. Additional sequences in the clone: 'ACC' after SalI site  
and before 'ATG' to provide Kozak consensus sequence; 'GG' after  
last codon and before HindIII site to maintain reading frame.  
Clone distribution: <http://bioinfo.clontech.com/orfclones>.

## FEATURES

source

Location/Qualifiers

1..882  
/organism="synthetic construct"  
/mol\_type="mRNA"  
/db\_xref="taxon:32630"  
/clone="GH00284L1.0"  
/clone\_lib="BD Creator(TM) CDS Library derived from MGC  
collection"  
/lab\_host="DH5alpha T1 resistant"  
/note="Vector: pDNR-Dual"  
1..->882  
/note="Mutations: 881:Stop->Leu"  
/codon\_start=1  
/transl\_table=1  
/product="Homo sapiens kallikrein 5"  
/protein\_id="AAP36503.1"  
/db\_xref="GI:30584501"

## CDS

1..->882  
/note="Mutations: 881:Stop->Leu"  
/codon\_start=1  
/transl\_table=1  
/product="Homo sapiens kallikrein 5"  
/protein\_id="AAP36503.1"  
/db\_xref="GI:30584501"  
GNSQDLGAGEDARSDSSRIINGSCDHMTQWQAALLRLPNQLVCGAVLHPQW  
LTAARCKRVFRVRLGHYSUSPVYSGQMFQGVKSIPHPGYSHPGHNDLMLIKLN  
RRIRPTKRVINVSHPGTSAGTKCLVSGWGTTKSPQVHPFKVLQCLNIVLSQRKE  
DAYPRQIDTFMFCAGDKAGSCQDGGVYVCGSLQGLVSWGIDYPCARPNRPGVYT  
NLCKFTWIKETIQANSL"

BASE COUNT 192 a 272 c 237 g 181 t  
ORIGIN

## Alignment Scores:

Pred. No.: 8,11e-121 Length: 882  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 12 Gaps: 0

US-09-936-271B-14 (1-293) x BT007831 (1-882)

QY 1 MetaAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
Db 1 ATGGGTACAGCAAGACCCCTGGATGTGGGTCTGTGCTCTGATCACACCTTGCTT 60  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
Db 61 CTGGGGGTTCACAGACATGTTCTCCCAACAATGATGTTCTGTGACCCCTCTAC 120  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyCysAlaArgSer 60  
Db 121 ACCGTGCTCTCGGAGCAACACGAGCTGGGAGCTGGGGCCGGGGAAGACCCCGTGC 180  
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80  
Db 181 GATGACAGCAGCGCGCATCATCAATGATCGGATGCGATGATGATGATGATGATGAT 240  
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuHisPro 100  
Db 241 CAGGCGCGCTGCTTCTAAGGCCCAACACGCTCTACTGCGGGCGGTGTGTGTGATCA 300  
QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120

Db 301 CAGTGGCTGCTCAGCGCCGCCACCTGCAGAGAAAGTTTTCAGAGTCCGCTCTCGGCCAC 360  
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
Db 361 TACTCCCTGTCCACAGTTTATGAATGGCGACAGATGTCAGGGGTCAATCATC 420  
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
Db 421 CCCCACCTGGCTACTCCCACTGACCTCTAAGGACCTCATGCTCATCAACTGAAC 480  
QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
Db 481 AGAAGAATTCGTCCTCCACTAAGATGTCAGACCATCAAGCTCTCTCTCATTCGCTCT 540  
QY 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200  
Db 541 GCTGGACAAAGTGTGTGTGCTGGTGGGGACAACCAAGAGCCCCCAAGTGCATCTC 600  
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
Db 601 CCTAAGGTCTCCAGTGTCTGAATATCAGCGTCTAAGTCAGAAAAGTGCAGAGGTCT 660  
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
Db 661 TACCGGAGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAAGCAGTAGAGACTCC 720  
QY 241 CysGlnGlyAspSerGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
Db 721 TCCAGGGGTGATTCGGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGCTCGTGTCC 780  
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
Db 781 TGGGGAGATTACCTGTGTCGGCCGACACAGACCGGGTGTCTACACGAACTCTGCAAG 840  
QY 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
Db 841 TTCACCAAGTGGATCCAGGAAACCATCCAGGCAACTCC 879

RESULT 3  
AY279380 1370 bp mRNA linear PRI 26-MAY-2003  
Homo sapiens kallikrein 5 splice variant 1 (KLK5) mRNA, complete  
cds; alternatively spliced.  
AY279380  
AY279380.1 GI:31075480  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1370)  
AUTHORS Kurlender,L., Yousef,G.M., White,N.M.A., Robb,J.-D., Borgono,C.A.  
and Diamandis,E.P.  
TITLE Identification of splice variants for the human kallikrein gene 5  
(KLK5)  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 1370)  
Kurlender,L., Yousef,G.M., White,N.M.A., Robb,J.-D., Borgono,C.A.  
and Diamandis,E.P.  
TITLE Direct Submission  
JOURNAL Submitted (19-APR-2003) Pathology and Laboratory Medicine, Mount  
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,  
Canada  
FEATURES Location/Qualifiers  
source 1..1370  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/map="19q13.4"  
1..1370  
/gene="KLK5"

```

5'UTR      1. .127
/gene="KLK5"
CDS        128. .1009
/gene="KLK5"
/notes="serine protease; hk5; alternatively spliced;
similar to the product of GenBank Accession Number
AF135028"
/codon_start=1
/product="kallikrein 5 splice variant 1"
/protein_id="AAP42275.1"
/db_xref="GI:31075481"
/translation="MATARPWWVLCALITALLLGVTEHLVLANNDVSCDHPNTVPS
GSNDLGAGAGEDARDSSRLINGSDDMTQPWQALLLPNQLYCAVALVHPQW
LTPAARCKKVFVRGLGHYSLSPEYSGQQMFQVKSIPHPGYSHPGHNDMLIKLN
RRLRPTKDVRIPIINVSHCPISAGTKCLVSGWGTQSPQVHPFKVLQCLNTSVLSQKRC
DAYPROIDDTMFCAGDKAGRDSQGGSPVVCNGSLQGLVSGWDYPCARPNRPGVYT
NLCKFTKWIQETIQANS"
1010. .1370
/gene="KLK5"
polyA_signal 1319. .1324
/gene="KLK5"
BASE COUNT 323 a 424 c 352 g 271 t
ORIGIN

Alignment Scores:
Pred. No.: 1.3e-120 Length: 1370
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-936-271B-14 (1-293) x AY279380 (1-1370)

Qy 1 MetAlaThrAlaArgProTrrPmetTrpValLeuCysAlaLeuIleThrAlaLeu 20
Db 128 ATGCTATACACAGACCCCTGGATGGTGGTCTCTGCTCTGATCAGAGCTTGCTT 187

Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 188 CTGGGGGTACAGACATGTTCTCGCCCAACAATGATGTTTCTGTGACCACTCTTAAC 247

Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db 248 ACCGTGCCCCCTGGGAGCAACAGACCTGGGAGCTGGGGCCGGGGAAGACGCCCGCTG 307

Qy 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrrP 80
Db 308 GATGACAGCAGCAGCCCATCATCATGATGGATCCGACTGCCATATGCACACCCAGCCGTG 367

Qy 81 GlnAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 368 CAGGCCCGCTGTTGCTAAGGCCCAACAGCTCTACTCGGGCGGGTGGTGGTGCATCCA 427

Qy 101 GlnTrpLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
Db 428 CAGTGGCTGTACGGCCGCCACTCGAGGAAGAAGTTTTCAGAGTCCGCTCGGCCAC 487

Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 488 TACTCCCTGTACACAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGGTCAATCCATC 547

Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
Db 548 CCCACCCCTGGCTACCCACCTGGCCACTCTAACGACCTCATGCTCATCAACTGAAC 607

Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db 608 AGAAGATTCTGCTCCCACTAAAGATGTGAGACCCATCAACGCTCTCTCATTTGCTCCCT 667

Qy 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200
Db 668 GCTGGGCAAAAGTCTTGGTGTCTGGCTGGGGGACAAACCAAGAGCCCCCAAGTGCATTC 727

```

```

Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db 728 CCTAAGGTCCTCCAGTCTTGAATATCAGCGTGTCTAAGTCAGAAAGGTGCGAGGATGCT 787

Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db 788 TACCCGAGACAGATAGATGACACATGTTCTGGCGCGGTGACAAAGCAGGTAGAGACTCC 847

Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db 848 TGCAGGGTGATTTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTGCTGCTCC 907

Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 908 TGGGGAGATTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAG 967

Qy 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db 968 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1006

RESULT 4
BD107879
LOCUS      BD107879      1381 bp      DNA      linear      PAT 18-SEP-2002
DEFINITION 36 human secreted proteins.
ACCESSION  BD107879
VERSION    BD107879.1 GI:23202697
KEYWORDS   JP 2002500035-A/50.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 1381)
AUTHORS    Ruben, S.M., Soppet, D.R., Ebner, R., Lafleur, D.W., Ni, J.,
            Brewer, L.A., Olsen, H.S., Duan, R.D. and Rosen, C.A.
TITLE      36 human secreted proteins
JOURNAL    Patent: JP 2002500035-A 50 08-JAN-2002;
            HUMAN GENOME SCIENCES INC
COMMENT     OS Homo sapiens (human)
            PN JP 2002500035-A/50
            PD 08-JAN-2002
            PF 06-JAN-1999 JP 2000527554
            PR 07-JAN-1998 US 60/070657, 07-JAN-1998 US 60/070658 PR
            07-JAN-1998 US 60/070692, 07-JAN-1998 US 60/070704 PI STEVEN
            NI RUBEN, DANIEL R SOPPET, REINHARD EBNER, DAVID W LAFLEUR, PI
            JIAN
            PI LAURIE A BREWER, HENRIK S OLSEN, ROSANNE D DUAN, CRAIG A ROSEN PC
            C12N15/09, A61K31/711, A61K38/00, A61K39/395, A61K39/395, A61K48/00,
            PC A61P5/00, A61P11/06, A61P19/02, A61P29/00, A61P31/18, A61P35/00, PC
            A61P35/02,
            PC A61P37/00, C07K14/435, C07K16/18, C12N1/15, C12N1/21, C12N5/10, PC
            C12N15/00,
            PC A61K37/02, C12N5/00
            CC 36 human secreted proteins
            FH Key Location/Qualifiers
            FT source 1..1381
            /organism="Homo sapiens (human)"
            /db_xref="taxon:9606"

BASE COUNT 329 a 425 c 353 g 274 t
ORIGIN

Alignment Scores:
Pred. No.: 1.31e-120 Length: 1381
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

```

US-09-936-271B-14 (1-293) x BD107879 (1-1381)	
QY	1 MetalAlaThrAlaArgProProTrrpMetTrrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
Db	129 ATGGCTACAGCAAGACCCCGCTGGATGTGGGTGCTCTGTCTGCTATCATCAGCGCTTGCTT 188
QY	21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db	189 CTGGGGTCACAGAGCATGTTCTGCCCAACATGATGTTCTCTGTGACCACTCTTAAC 248
QY	41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db	249 ACCGTGCCCTCTGGGAGCAACACAGGACCTGGAGCTGGGGCCGGGGAAGACGCCCGCTCG 308
QY	61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrrp 80
Db	309 GATGACAGCAGACGCCGCATCATCAATGGATCGGACTGCGATGACACACCCACCGCTGG 368
QY	81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrrpCysGlyAlaValLeuValHisPro 100
Db	369 CAGGCGCGCTGTGTAGGCCCAACACAGCTACTGCGGGCGGTGTGTGTCATCCA 428
QY	101 GlnTrrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
Db	429 CAGTGGCTGCTACGGCCGCCACTGCAGGAAGAAAGTTTTCAGAGTCCGTCTCGGCCAC 488
QY	121 TyrSerLeuSerProValTrrpGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db	489 TACTCCCTGTACCAAGTTTATGAATCTGGGACGACATGTTCCAGGGGTCAATCCATC 548
QY	141 ProHisProGlyTrrpSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
Db	549 CCCCACCTGGCTACTCCACCTGGCCACTTAAAGACCTCATGCTCATCAAACTGAAC 608
QY	161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db	609 AGAAGAATTCTGCTCCCACTAAAGATGTACAGACCATCAAGCTCTCTCTCATGTCCTCT 668
QY	181 AlaGlyThrLysCysLeuValSerGlyTrrpGlyThrThrLysSerProGlnValHisPhe 200
Db	669 GCTGGGCAAGAGTCTGTGTGCTGTGGCTGGGGGACCAACCAAGAGCCCCCAAGTGCATTC 728
QY	201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db	729 CCTAAGTCTCTCCAGTGCTGAATATCAGGTGCTAAGTCAGAAAAGGTGCGAGGATGCT 788
QY	221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db	789 TACCCGACAGACATAGATACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCC 848
QY	241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db	849 TGGCAGGGTATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 908
QY	261 TrrpGlyAspTrrpProCysAlaArgProAsnArgProGlyValTrrpThrAsnLeuCysLys 280
Db	909 TGGGGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACTCTCCAG 968
QY	281 PheThrLysTrrpIleGlnGluThrIleGlnAlaAsnSer 293
Db	969 TTCACCAAGTGTATCCAGGAACCACTCCAGGCCAATCC 1007
RESULT 5	
AF168768	
LOCUS	Homo sapiens stratum corneum trypsin-like serine (SCTE) mRNA, complete cds.
DEFINITION	
ACCESSION	AF168768
VERSION	AF168768.2
KEYWORDS	GI:20153423
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
1 (bases 1 to 1387)	
Brattsand, M. and Egelrud, T.	
Purification, molecular cloning, and expression of a human stratum corneum trypsin-like serine protease with possible function in desquamation	
J. Biol. Chem. 274 (42), 30033-30040 (1999)	
99445563	
10514489	
2 (bases 1 to 1387)	
Brattsand, M. and Egelrud, T.	
Submitted (12-JUL-1999) Public Health and Clinical Medicine, Dermatology and Venereology, Umea University, University Hospital, Umea SE-901 85, Sweden	
3 (bases 1 to 1387)	
Brattsand, M. and Egelrud, T.	
Submitted (16-APR-2002) Public Health and Clinical Medicine, Dermatology and Venereology, Umea University, University Hospital, Umea SE-901 85, Sweden	
Sequence update by submitter	
On Apr 16, 2002 this sequence version replaced gi:6063032.	
FEATURES	
Location/Qualifiers	
1..1387	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/tissue_type="stratum corneum"	
1..1387	
/gene="SCTE"	
170..1051	
/gene="SCTE"	
/note="SCTE; trypsin-like serine protease; possibly involved in desquamation"	
/codon_start=1	
/product="stratum corneum tryptic enzyme"	
/protein_id="AAF03101.1"	
/db_xref="GI:6063033"	
/translation="MATARPPMWMVLCALITALLGVTEHVLANNVSCDHPNTVPVS GSNQDIGAGEDARDSSSRITNGSDCMHTQPWQAALLRPNGLYCGAVLVHPQM LITAAHCRKKVFRVRLGHYSLSPVYESGQMGQVKSIPHPGYSHFGHSNDMLIKLN RRRPRTKDVRIENVSSHCPAGTKLVSGWGTTPKSPQVHPFKVLOCLNTSVLSOKRCE DAYPRQIDDTMFCAGDKGRDSCQSGSGGPPVVCNGSLQGLVSGDYPICARPNRGVYT NLCKTKWIQETIQANS"	
BASE COUNT 290 a 437 c 351 g 309 t	
ORIGIN	
Alignment Scores:	
Pred. No.: 1..31e-120 Length: 1387	
Score: 1608.00 Matches: 293	
Percent Similarity: 100.00% Conservative: 0	
Best Local Similarity: 100.00% Mismatches: 0	
Query Match: 100.00% Indels: 0	
DB: 9 Gaps: 0	
US-09-936-271B-14 (1-293) x AF168768 (1-1387)	
QY	1 MetalAlaThrAlaArgProProTrrpMetTrrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
Db	170 ATGGCTACAGCAAGACCCCGCTGGATGTGGGTGCTCTGTCTGCTATCATCAGCGCTTGCTT 229
QY	21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db	230 CTGGGGTCACAGAGCATGTTCTGCCCAACATGATGTTCTCTGTGACCACTCTTAAC 289
QY	41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db	290 ACCGTGCCCTCTGGGAGCAACACAGGACCTGGGAGCTGGGGCCGGGGAAGACGCCCGCTCG 349
QY	61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrrp 80

```

Db 350 GATGACAGCAGCAGCCGCATCATCAATGGATCGGACTCGCATATGCGACACCCAGCCGCTGG 409
Qy 81 GlnAlaAlaLeuLeuLeuAraGProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 410 CAGCCCGCGCTGTGCTAAGCCCAACACAGCTCTACTGCGGGCGGTGTGGTGCATCCA 469
Qy 101 GlnTrpLeuLeuThraAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
Db 470 CAGTGGCTGCTACGCGCGCCCACTGCGAGGAAGAAGTTTTCAGAGTCCGCTCGCGCCAC 529
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 530 TACTCCCTGTACACAGTTATGAATCTGGCAGCAGATGTTCAGGGGGTCAATCCATC 589
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160
Db 590 CCCACACCTGGCTACTCCACCTGGCCACTCTAACGACCTCATGCTCATCAACTGAAC 649
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db 650 AGAAGAATTCTGCCACTAAAGATGTCCAGACCCATCAACGTCTCTCTATTGTCCCTCT 709
Qy 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200
Db 710 GCTGGGACAAAGTCTGGTGTCTGGCTGGGGGACCAACCAAGACCCCAAGTGCACATC 769
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db 770 CCTAAGTCTCTCCAGTCTTGAATATCAGCGTGTAGTCAGAAAGAGTGGCAGGATGCT 829
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db 830 TACCCGAGACAGATAGATGACACCATCTTCTGCGCGGTGCACAAAGCAGGTAGAGACTCC 889
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db 890 TGCAGGGGTGATTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 949
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 950 TGGGAGATFACCTTGTGCGCGGCCCAACAGACCCGGGTGTCTACACGACCTCTGCAAG 1009
Qy 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db 1010 TTCACCAAGTGATCCAGGAACCATCCAGGCCAACTCC 1048

RESULT 6
LOCUS AY279381 1438 bp mRNA linear PRI 26-MAY-2003
DEFINITION Homo sapiens kallikrein 5 splice variant 2 (KLK5) mRNA, complete
cds; alternatively spliced.
ACCESSION AY279381
VERSION AY279381.1 GI:31075482
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1438)
AUTHORS Kurlender,L., Yousef,G.M., White,N.M.A., Robb,J.-D., Borgono,C.A.
and Diamandis,E.P.
TITLE Identification of splice variants for the human kallikrein gene 5
(JKLK5)
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1438)
AUTHORS Yousef,G.M., Robb,J.-D., White,N.M.A., Kurlender,L., Borgono,C.A.
and Diamandis,E.P.
TITLE Direct Submission
JOURNAL Submitted (19-APR-2003) Pathology and Laboratory Medicine, Mount
Sinai Hospital, 600 University Avenue, Toronto, Ontario M5G 1X5,
Canada
FEATURES
source Location/Qualifiers
1. .1438

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19q13.4"
1. .1438
/gene="KLK5"
1. .195
/gene="KLK5"
196. .1077
/gene="KLK5"
/notes="serine protease; differentially expressed in
malignancy; may play a role in desquamation of skin; hK5;
alternatively spliced; similar to the product of GenBank
Accession Number AF135028"
/codon_start=1
/product="kallikrein 5 splice variant 2"
/protein_id="AAP42276.1"
/db_xref="GI:31075483"
/translation="MATARPMMWVLCALITALLGVTEHVLANNNDVSCDHPNSTVPS
GSNODLGAGAGEDARSDSSRIINGSQDMHTOPWQAALLRPNOYLCGAVLVHPQW
LLTAACHCKKVFVRVLGHYSISPVESQDMFOGVKSIHPGVSHPGHNSDLMLIKLN
RRIRPKDVRPINVSSHCPAGTKLVSGHGTTSKSPQVHPFKVLQCLNLSVLSQKRC
DAYPRQIDDTMFCAGDKAGRDSQGDSSGPPVNCVGLSVGWDYPCARNRPGVYT
NLCKFTKWIQETIQANS"
1078. .1438
/gene="KLK5"
1387. .1392
/gene="KLK5"
332 a 443 c 374 g 289 t
BASE COUNT
ORIGIN
Alignment Scores:
Pred No.: 1.37e-120 Length: 1438
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0
US-09-936-271B-14 (1-293) x AY279381 (1-1438)
Qy 1 MetaThraAlaArgProTrpMetTrpValLeuCysAlaLeuIleThraAlaLeuLeu 20
Db 196 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTCTGTCTGTATGCACAGCCTTGTCT 255
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn* 40
Db 256 CTGGGGGTACAGACGATGTTCTGCCAACAATGATGTTCTCTGACCCCTCTAAC 315
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db 316 ACCGTGCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCGGGAGAGCGCCGCTCG 375
Qy 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 376 GATCAGCAGCAGCAGCCGATCATCAATGATGATCCGACTGCGATATGACACCCAGCCGCTGG 435
Qy 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 436 CAGGCGCGCTGTGTGCTAAGGCCCAACAGCTCTACTGCGGGCGGTGTGGTGCATCCA 495
Qy 101 GlnTrpLeuLeuThraAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
Db 496 CAGTGGTGTCTACGGCGCCCACTGCGAGGAAGAAGTTTTCAGAGTCCGCTCGCGCCAC 555
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 556 TACTCCCTGTACACAGTTATGAATCTGGGCGAGCAGATGTTCAGGGGGTCAATCCATC 615
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160
Db 616 CCCACACCTGGCTACTCCACCCCTGGCCACTCTAACGACCTCATGCTCATCAACTGAAC 675

```





/clone="MGC:1107 IMAGE:2989806"  
/tissue\_type="Ovary, adenocarcinoma"  
/clone\_lib="NIH\_MGC\_9"  
/lab\_host="DH10B-R"

/note="vector: pOTB7"

CDS

294..1175

/codon\_start=1

/product="kallikrein 5"

/protein\_id="AAH08036.1"

/db\_xref="GI:14165478"

/translation="WATARPWWVLCALITALLGLGVTEHVLANNVDSCHDPSNTVPS  
GNSQDAGAGEDARSDSSRIINGSCDMHTQWQAALLRLPNQLYCGAVLVHPQW  
LITAAHCRKRVRLGHYSLSPVYESQGMFQGVKSIPIPGYSHPGHNDLMLIKLN  
RRIRPTKVRPNVSSHCPAGTKICLVSGWGTTPQVHFPRKVLQCLNLSVLSQRCE  
DAYPRIDDTMFCAGDKAGRDSCQGDGSGPVVCGSLQGLVSGWDYPCARPNRPGVYT  
NLCKFTKWIQETIQANS"

BASE COUNT 346 a 467 c 401 g 322 t

ORIGIN

Alignment Scores:

Pred. No.: 1,47e-120 Length: 1536  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 9 Gaps: 0

US-09-936-271B-14 (1-293) x BC008036 (1-1536)

QY 1 MetAlaThraAlaArgProProtrpMetTrpValLeuCysAlaLeuLeuLeuLeuLeu 20  
DB 294 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTGCTGTGATCAGACCTTGCTT 353  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
DB 354 CTGGGGGTACAGAGCATGTTCTGCCAACAATGATGTTCTGTGACCACTTCTAAC 413  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyAlaArgSer 60  
DB 414 ACCGTGCCCTCTGGAGAACACAGGACCTGGGAGCTGGGCGCGGGAACGCCGTG 473  
QY 61 AspAspSerSerArgIleLeuAsnGlySerAspCysAspMethHisThrGlnProtrp 80  
DB 474 GATGACAGCAGCAGCGCATCATCAATGATGATGATGATGATGATGATGATGATG 533  
QY 81 GlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100  
DB 534 CAGGCGCGCTGTGTAAAGGCCCAACACAGCTTACTGCGGGCGGTGTGTGTCATCCA 593  
QY 101 GlnTrpLeuLeuThraAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120  
DB 594 CAGTGGCTGCTCAGCGCGCCCACTCCAGGAGAAAGTTTTCAGAGTCCGTCTCGGCC 653  
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
DB 654 TACTCCCTGTACACAGTTTATGAATCTGGGACAGCATGTTCCAGGGGTCAATCATC 713  
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeu 160  
DB 714 CCCACCCCTGGTACTCCACCCCTGGCCACTCTAAGAGCTCATGCTCATCAACATG 773  
QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
DB 774 AGAAGAAATTCGTCCTAAAGATGTCAGACCATCAAGCTCTCTCTCTCTCTCTCT 833  
QY 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
DB 834 GCTGGGACAAAGTCTGTGTGCTGGTGGGGACCAACCAAGAGCCCAAGTGCATTC 893  
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
DB 894 CCTAAGTCTCTCAGTGTGTAATATCATCAGCGTCTAAGTCAGAAAAAGGTGCGAGG 953

QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
DB 954 TACCCGAGACAGATAGATGACACCATGTTCTGGCGGGTGACAAACAGGTAGACTCC 1013  
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
DB 1014 TGCCAGGGTGATTCTGGGGGGCTGTGTGCTGCAATGGCTCCCTGAGGAGCTCGTGTCC 1073  
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
DB 1074 TGGGGAGATTAACCTTGTGCCCGCCCAACAGCCGGGTGTCTACACGAACCTCTCAAG 1133  
QY 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
DB 1134 TTCACCACTGGATCCAGGAACCATCCAGGCCAATCTCC 1172  
RESULT 10  
AR252580  
LOCUS AR252580 1570 bp DNA linear PAT 20-DEC-2002  
DEFINITION Sequence 308 from patent US 6478825.  
ACCESSION AR252580  
VERSION AR252580.1 GI:27300488  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1570)  
AUTHORS Winterbottom,J.M., Shimp,L., Boyce,T.M. and Kaes,D.  
TITLE Implant, method of making same and use of the implant for the  
treatment of bone defects  
JOURNAL Patent: US 6478825-A 308 12-NOV-2002;  
FEATURES Location/Qualifiers  
source 1..1570  
/organism="unknown"  
BASE COUNT 332 a 485 c 424 g 329 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,5e-120 Length: 1570  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 6 Gaps: 0  
US-09-936-271B-14 (1-293) x AR252580 (1-1570)  
QY 1 MetAlaThraAlaArgProProtrpMetTrpValLeuCysAlaLeuLeuLeuLeuLeu 20  
DB 354 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTGCTGTGATCAGACCTTGCTT 413  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
DB 414 CTGGGGGTACAGAGCATGTTCTGCCAACAATGATGATGATGATGATGATGATGATG 473  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyAlaArgSer 60  
DB 474 ACCGTGCCCTCTGGGAGAACACAGGACCTGGGAGCTGGGCGCGGGAACGCCGTG 533  
QY 61 AspAspSerSerArgIleLeuAsnGlySerAspCysAspMethHisThrGlnProtrp 80  
DB 534 GATGACAGCAGCAGCGCATCATCAATGATGATGATGATGATGATGATGATGATG 593  
QY 81 GlnAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100  
DB 594 CAGGCGCGCTGTGTAAAGGCCCAACACAGCTTACTGCGGGCGGTGTGTGTCATCCA 653  
QY 101 GlnTrpLeuLeuThraAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120  
DB 654 CAGTGGCTGCTCAGCGCGCCCACTCCAGGAGAAAGTTTTCAGAGTCCGTCTCGGCC 713  
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140

```
Db 714 TACTCCCTGTCACACAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGGTCAAAATCCATC 773
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuHisLeuAsn 160
Db 774 CCCACCTGGCTACTCCACCTCGGCACCTCTAACGACCTCATGCTCATCAACTGAAC 833
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db 834 AGAAGAATTCGTCCTCCACTAAAGATGTCCAGACCCATCAACGCTCCTCTCATTTGCCCTCT 893
Qy 181 AlacGlyThrLysCysLeuValSerGlyTyrGlyThrLysSerProGlnValHisPhe 200
Db 894 GCTGGGACAAGTCTGGTGTCTGGCTGGGGGACAAACCAAGACCCCAAGTGCATTC 953
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db 954 CCTAAGGTCTCCAGTCTTGAATATCAGCGTGTCTAAGTCAGAAAGTGGAGGATGCT 1013
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db 1014 TACCCGAGACAGATAGATGACACCATGTTCTGGCCGGTGACAAAGCAGGTAGAGACTCC 1073
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyValSer 260
Db 1074 TGCAGGGTGATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAAGGACATCGTGCTC 1133
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 1134 TGGGAGATTTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAG 1193
Qy 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db 1194 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1232

RESULT 11
LOCUS AX080829 1570 bp DNA linear PAT 27-FEB-2001
DEFINITION Sequence 75 from Patent WO0109327.
ACCESSION AX080829
VERSION AX080829.1 GI:13169796
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
AUTHORS Ashkenazi,A.J., Baker,K.P., Goddard,A., Godowski,P.J., Gurney,A.L.,
KlJavin,I.J., Lafleur,M., Mark,M.R., Marsters,S.A., Pitti,R.M.,
Watanabe,C.K. and Wood,W.I.
TITLE Method of preventing the injury or death of retinal cells and
treating ocular diseases
JOURNAL Patent: WO 0109327-A 75 08-FEB-2001;
Genentech, Inc. (US)
FEATURES
source
1..1570
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
BASE COUNT 332 a 485 c 424 g 329 t
ORIGIN
Alignment Scores:
Pred. No.: 1.5e-120 Length: 1570
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-936-271b-14 (1-293) x AX080829 (1-1570)
Qy 1 MetAlaThrAlaArgProTyrPheMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
|||||
```

```
Db 354 ATGGCTACACAAAGACCCCTCGATGTGGGTGCTCTGTGCTGTGATCATCAGACCTTGCTT 413
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
|||||
Db 414 CTGGGGTTCACAGACATGTTCTCGCAACAATGATGTTCTTGACACACCTCTTAAC 473
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
|||||
Db 474 ACCGTGCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCGGGGAAGACGCCGGTCG 533
Qy 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMethHisThrGlnProTrp 80
|||||
Db 534 GATGACAGCAGCAGCGCATCATCAATGATGCGACTGCGCATATGACACCCAGCCGTGG 593
Qy 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
|||||
Db 594 CAGGCCGCGCTGTTGCTAAGGCCCAACGACCTCTACTGCGGGGGTGTGGTGCATCCA 653
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120
|||||
Db 654 CAGTGGGTGCTCAGCGCGCCACCTGCAGGAAGAAAGTTTTCAGAGTCCGCTCTCGGCCAC 713
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
|||||
Db 714 TACTCCTGTCCACAGTTTATGAATCTGGCAGCAGATGTCCAGGGGGTCAAAATCCATC 773
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160
|||||
Db 774 CCCACCTGGCTACTCCACCTGGCCACTCTAACGACCTCATGCTCATCAAACTGAAC 833
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
|||||
Db 834 AGAAGAATTCGTCCTCCACTAAAGATGTCCAGACCCATCAACGCTCTCTCTATTGCTCTCT 893
Qy 181 AlacGlyThrLysCysLeuValSerGlyTyrGlyThrLysSerProGlnValHisPhe 200
|||||
Db 894 GCTGGGACAAGTGTGGTGTCTGGCTGGGGGACAAACCAAGACCCCAAGTGCATTC 953
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db 954 CCTAAGGTCTCCAGTCTTGAATATCAGCGTGTCTAAGTCAGAAAGTGGAGGATGCT 1013
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
|||||
Db 1014 TACCCGAGACAGATAGATGACACCATGTTCTGGCCGGTGACAAAGCAGGTAGAGACTCC 1073
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyValSer 260
Db 1074 TGCAGGGTGATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAAGGACATCGTGCTC 1133
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 1134 TGGGGAGATTTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAG 1193
Qy 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
|||||
Db 1194 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1232

RESULT 12
AX03421 1570 bp DNA linear PAT 14-JUN-2002
LOCUS AX03421
DEFINITION Sequence 308 from Patent WO0073454.
ACCESSION AX03421
VERSION AX03421.1 GI:21436934
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1
AUTHORS Ashkenazi,A.J., Baker,K.P., Botstein,D., Desnoyers,L., Eaton,D.,
Ferrara,N., Gerber,H., Gerritsen,M., Goddard,A., Godowski,P.,
Grimaldi,C.J., Gurney,A.L., KlJavin,I., Napier,M.A., Pan,J.,
```

Paoni,N.F., Roy,M., Stewart,T.A., Tumas,D., Watanabe,C.K.,  
Williams,P., Wood,W.I. and Zhang,Z.  
Secreted and transmembrane polypeptides and nucleic acids encoding  
the same

JOURNAL Patent: WO 0073454-A 308 07-DEC-2000;

Genentech Inc. (US)

FEATURES Location/Qualifiers

source 1..1570

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 332 a 485 c 424 g 329 t

ORIGIN

Alignment Scores:

Pred. No.: 1..1570 Length: 1570

Score: 1608.00 Matches: 293

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-936-271B-14 (1-293) x AX403421 (1-1570)

QY 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
DB 354 ATGGCTACAGCAAGACCCCTCGATGGTGCTCTGTCTGATCAGCGCTTGCCT 413  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
DB 414 CTGGGGGTACAGAGCATGTTCTCGCCACAATGATGTTCTGTGACCACTCTTAAC 473  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60  
DB 474 ACCGTGCCCTCTGGGACACACAGACCTGGGAGCTGGGCGGGGAAGCGCCGCTG 533  
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMethIleThrGlnProTrp 80  
DB 534 GATGACAGCAGCAGCGCATCATCAATGGATCGGACTGCGATATGCACACCCGCTGG 593  
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrpCysGlyAlaValLeuValHisPro 100  
DB 594 CAGCGCGCGCTGTGTAAAGCCCAACAGCTCTACTGCGGGCGGTGTGTGTGATCCA 653  
QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120  
DB 654 CAGTGGCTGCTCAGCGCGCCCACTGCAGGAAGAAAGTTTCAGAGTCCGCTCGGCCAC 713  
QY 121 TyrSerLeuSerProValTrpGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
DB 714 TACTCCCTGTCCAGCTTTATGATCTGGGCACAGATGTTCCAGGGGTCAATCCATC 773  
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
DB 774 CCCACCCCTGGCTACCCACCTCGCCACTCTAAGACCTCATGCTCATCAACTGAAC 833  
QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
DB 834 AGAAGAAATTCGTCCTCCCAATAAGATGTCAGCCCATCAACGTCCTCTCATTTGCTCT 893  
QY 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
DB 894 GCTGGGACAAAGTCTGTGTGCTGGCTGGGGACAACCAAGAGCCCCCAAGTGCACTTC 953  
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
DB 954 CCTAAGTCTCTCCAGTCTTGAATATCAGCGTCTTAAGTCAGAAAAAGGTGCGAGGATGCT 1013  
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
DB 1014 TACCCGACAGATAGATACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCC 1073  
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260

DB 1074 TGCAGGGTCATCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTCAGGAGCTGTGTCC 1133  
QY 261 TrpGlyAspTrpProCysAlaArgProAsnArgProGlyValTrpThrAsnLeuCysLys 280  
DB 1134 TGGGGAGATTACCTTGTGCGCGCCCAACAGACCGGGTGTCTACACGAACCTCTCAAG 1193  
QY 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
DB 1194 TTCACCAAGTGGATCCAGGNAACCATCCAGGCCAATCC 1232

RESULT 13

AX464322

LOCUS

DEFINITION

AX464322

ACCESSION

VERSION

AX464322.1

GI:21899177

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

(human)

REFERENCE

AUTHORS

Baker,K.P., Beresini,M., Deforge,L., Desnoyers,L., Filvaroff,E.,

Gao,W.Q., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,

Sherwood,S., Smith,V., Stewart,T.A., Tumas,D., Watanabe,C.K.,

Wood,W.L. and Zhang,Z.

Secreted and transmembrane polypeptides and nucleic acids encoding

same

JOURNAL

Patent: WO 0140466-A 455 07-JUN-2001;

Genentech Inc. (US)

FEATURES

source

1..1570

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

BASE COUNT 332 a 485 c 424 g 329 t

ORIGIN

Alignment Scores:

Pred. No.: 1..1570 Length: 1570

Score: 1608.00 Matches: 293

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 100.00% Indels: 0

DB: 6 Gaps: 0

US-09-936-271B-14 (1-293) x AX464322 (1-1570)

QY 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
DB 354 ATGGCTACAGCAAGACCCCTCGATGGTGCTCTGTCTGATCAGCGCTTGCCT 413  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
DB 414 CTGGGGGTACAGAGCATGTTCTCGCCACAATGATGTTCTGTGACCACTCTTAAC 473  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60  
DB 474 ACCGTGCCCTCTGGGACACACAGACCTGGGAGCTGGGCGGGGAAGCGCCGCTG 533  
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMethIleThrGlnProTrp 80  
DB 534 GATGACAGCAGCAGCGCATCATCAATGGATCCGACTGCGATATGCACACCCGCTGG 593  
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrpCysGlyAlaValLeuValHisPro 100  
DB 594 CAGCGCGCGCTGTGTAAAGCCCAACAGCTCTACTGCGGGCGGTGTGTGTGATCCA 653  
QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120  
DB 654 CAGTGGCTGCTCAGCGCGCCCACTGCAGGAAGAAAGTTTCAGAGTCCGCTCGGCCAC 713

QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
 DB 714 TACTCCCTGTCCACAGTTTGAATCTGGCAGCAGATGTTCCAGGGGTCAAAATCCATC 773  
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
 DB 774 CCCACACCTGGCTACTCCCACTGGCCACTTAACGACCATCAACGCTCATCAAACTGAAC 833  
 QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
 DB 834 AGAAGATTCTGCTCCCACTAAAGATGTCAGACCCATCAACGCTCTCTCTCATTTGCCCTCT 893  
 QY 181 AlaGlyThrLysCysLeuValSerGlyTyrPGLyThrThrLysSerProGlnValHisPhe 200  
 DB 894 GCTGGGACAAAGTCTGGTGTCTGGCTGGGGGACAAACCAAGACCCCAAGTGCACCTTC 953  
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
 DB 954 CCTAAGGTCTCTCCAGTCTTGAATATCAGGCTGCTAAGTCAGAAAGGTGCGAGGATGCT 1013  
 QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
 DB 1014 TACCCGACAGATAGATGACACCATCTTCTGCGCGGTGACAAAGCAGGTAGAGACTCC 1073  
 QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyValSer 260  
 DB 1074 TGCCAGGGTGATTCCTGGGGGCGCTGTGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 1133  
 QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
 DB 1134 TGGGAGATTACCTTGTGCGCGGCCCAACAGACCGGGGTCTACAGAACCTCTGCAAG 1193  
 QY 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
 DB 1194 TTCACCAAGTGATCCAGGAACCATCCAGGCCCAACTCC 1232  
 RESULT 14  
 LOCUS AR078184 1476 bp DNA linear PAT 31-AUG-2000  
 DEFINITION Sequence 2 from patent US 5962300.  
 ACCESSION AR078184  
 VERSION AR078184.1 GI:10004930  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1476)  
 AUTHORS Hillman, J. L. and Lal, P.  
 TITLE Human kallikrein  
 JOURNAL Patent: US 5962300-A 2 05-OCT-1999;  
 FEATURES Location/Qualifiers  
 source 1..1476  
 BASE COUNT 303 a 466 c 388 g 319 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,57e-119 Length: 1476  
 Score: 1595.00 Matches: 293  
 Percent Similarity: 99.66% Conservative: 0  
 Best Local Similarity: 99.66% Mismatches: 0  
 Query Match: 99.19% Indels: 1  
 DB: 6 Gaps: 0  
 US-09-936-271B-14 (1-293) x AR078184 (1-1476)  
 QY 1 MetalathrAlaArgProProThrMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
 DB 277 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTCTGTGATCAGACGCTTGTCT 336  
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
 DB 337 CTGGGGGTACAGAGATGTTCTCGCCCAACAAATGATTTCTCTGTGACCACCCCTCTAAC 396

QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
 DB 397 ACCGTGCCCTCTGGGAGCAACACGAGCCTGGGAGCTGGGCGGGGGAAGACGCCGCTCG 456  
 QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTyr 80  
 DB 457 GATGACAGCAGCAGCGCATCATATGGATCGACTGCGATATGACACCCAGCCGTGG 516  
 QY 81 -GlnAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPr 100  
 DB 517 CGAGGCGCGCTGTTGCTAAGGCCCAACAGCTCTACTGCGGGCGGTGTTGGTCATCC 576  
 QY 100 oGlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyH 120  
 DB 577 ACAGTGGCTGCTCACGGCGCCCACTGCAGGAAGAAAGTTTTCAGAGTCCGTCTCGGCCA 636  
 QY 120 sTyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIl 140  
 DB 637 CTACTCCCTGTCCAGTCTTATGAATCTGGGACAGATGTTCCAGGGGTCAAAATCCAT 696  
 QY 140 eProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAs 160  
 DB 697 CCCCACCTGGCTACTCCCACTGCGCCACTCTAACGACTCATGCTCATCAAACTGAA 756  
 QY 160 nArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSe 180  
 DB 757 CAGAAGAATTCGTCCTCCACTAAAGATGTCAGACCCATCAAGCTCTCTCTCATTTGCTCC 816  
 QY 180 rAlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPh 200  
 DB 817 TGCTGGGACAAAGTCTGGTGTCTGGTGTCCGGTGGGGGACAAACCAAGAGCCCCCAAGTGC 876  
 QY 200 eProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAl 220  
 DB 877 CCCTAAGGTCTCCAGTCTTGAATATCAGGTCTAAGTCAGAAAAAGTGCAGAGGATGC 936  
 QY 220 aTyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSe 240  
 DB 937 TTACCCGACAGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTC 996  
 QY 240 rCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSe 260  
 DB 997 CTGCCAGGGTGTCTCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTCGTGT 1056  
 QY 260 rTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLy 280  
 DB 1057 CTGGGGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTCTCTACAGAACCTCTCGAA 1116  
 QY 280 sPheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
 DB 1117 GTTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1156  
 RESULT 15  
 LOCUS AR137506 1476 bp DNA linear PAT 16-JUN-2001  
 DEFINITION Sequence 2 from patent US 6197511.  
 ACCESSION AR137506  
 VERSION AR137506.1 GI:14479015  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1476)  
 AUTHORS Hillman, J. L. and Lal, P.  
 TITLE Human kallikrein  
 JOURNAL Patent: US 6197511-A 2 06-MAR-2001;  
 FEATURES Location/Qualifiers  
 source 1..1476  
 BASE COUNT 303 a 466 c 388 g 319 t  
 ORIGIN

## Alignment Scores:

Pred. No.:	1-57e-119	Length:	1476
Score:	1595.00	Matches:	293
Percent Similarity:	99.66%	Conservative:	0
Best Local Similarity:	99.66%	Mismatches:	0
Query Match:	99.19%	Indels:	1
DB:	6	Gaps:	0

US-09-936-271b-14 (1-293) x AR137506 (1-1476)

QY	1	MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20
DB	277	ATGGCTACAGCAGACAGACCCCTGGATGTGGGTCTCTGTCTGTATCATCAGCCCTTGCTT 336
QY	21	LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
DB	337	CTGGGGGTACAGAGAGATGTTCTGCCAACACATGATGTTCTGTGTACCCCTCTTAAC 396
QY	41	ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
DB	397	ACCGTGCCTCTGGGAGCAACAGGACCTGGGAGCTGGGCGGGGCAAGACGCCGGTGC 456
QY	61	AspAspSerSerSerArgIleLeuAsnGlySerAspCysAspMethIsthrGlnProTrp 80
DB	457	GATGACAGCAGCAGCCGCATCATCAATGGATCCGACTGGCATATGCACACCCAGCCGTG 516
QY	81	-GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPr 100
DB	517	GCAGGCCGGCGTGTCTAAGGCCCAACAGCTCTACTCGGGGGCGGTGTGTGTGCTCC 576
QY	100	oGlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyH1 120
DB	577	ACAGTGGCTGCTCACGCCGCCACTGCAGGAAGAAAGTTTTCAGAGTCCGCTCTCGGCCA 636
QY	120	sTyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerI1 140
DB	637	CTACTCCCTGTACCACTTTATGAATCTGGGCAGCAGATGTTCCAGGGGGTCAAAATCCAT 696
QY	140	eProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAs 160
DB	697	CCCCACCTGGGTACTCCACCCTGGCCACTCTAACGACCTCATGCTCTCATCAAACTGAA 756
QY	160	nArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSe 180
DB	757	CAGAGAATTCGCCCACTAAAGATGTCAGACCCATCAACGCTCTCTCTCATTTGTCCTC 816
QY	180	rAlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPh 200
DB	817	TGCTGGGCAAAAGTGTGTGTGTCTCCGGCTGGGGGACAAACAGAGCCCCCAAGTGCAC 876
QY	200	eProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAl 220
DB	877	CCCTAAGTCTCTCCAGTGTGTGAATATCAGGTGCTTAAGTCAGAAAAGGTGCGAGGATC 936
QY	220	aTyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSe 240
DB	937	TTACCCGACAGACATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGACACTC 996
QY	240	rCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSe 260
DB	997	CTGCCAGGTGATTTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTG 1056
QY	260	rTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLy 280
DB	1057	CTGGGGAGATATACCTTTGTCGGGGCCCAACAGACCGGGTGTCTACCGAACTCTGCAA 1116
QY	280	sPheThrLysTrpIleGlnGlnThrIleGlnAlaAsnSer 293
DB	1117	GTTCAACAAGTGTGATCCAGAAACCATCCAGGCCCACTCC 1156

Search completed: August 5, 2003, 13:24:21  
Job time : 3136 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 5, 2003, 05:55:25 ; Search time 274 Seconds  
(without alignments)  
2886.625 Million cell updates/sec

Title: US-09-936-271B-14  
Perfect score: 1608  
Sequence: 1 MATARPPMVMVLCALITALL.....VYTNLCKTKTKWQETIQANS 293

Scoring table:  
BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0  
Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framet\_p2n.model -DEV=xlp  
-O=/cgn2\_1/USPTO.spool/US09936271/runat\_04082003\_104245\_25695/app\_query.fasta\_1.455  
-DB=N\_Geneseq\_19Jun03 -QFMT=fastp -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US09936271 -ECGN\_1\_1\_246 -runat\_04082003\_104245\_25695 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: /N\_Geneseq\_19Jun03:\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:\*  
25: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:\*  
26: /SIDSL/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed.

and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1608	100.0	882	20	AAZ23318 Human PDSP-1 DNA c
2	1608	100.0	1302	20	AAZ57989 Human BS247 specif
3	1608	100.0	1302	20	AAZ57990 Human BS247 specif
4	1608	100.0	1381	20	AAZ06259 Human secreted pro
5	1608	100.0	1499	20	AAZ60578 Human keratinocyte
6	1608	100.0	1516	20	AAZ06245 Human secreted pro
7	1608	100.0	1539	20	AAZ23317 Human PDSP-1 DNA.
8	1608	100.0	1570	21	AAZ65070 Membrane-bound pro
9	1608	100.0	1570	22	AAZ21471 Human cDNA sequenc
10	1608	100.0	1570	22	AAZ44216 Human PRO1132 (UNQ
11	1608	100.0	1570	24	ABK28605 Human DNA59767-148
12	1608	100.0	1570	25	ACA03830 cDNA encoding huma
13	1608	100.0	1570	25	ACA04251 Human cDNA encodin
14	1608	100.0	1570	25	ABX89368 DNA encoding novel
15	1608	100.0	1570	25	ABX80318 Human secreted or
16	1608	100.0	1570	25	ABX80822 Human secreted tra
17	1608	100.0	1570	25	ABX81205 Human secreted or
18	1608	100.0	1570	25	ABX90295 Human secreted tra
19	1608	100.0	1570	25	ABX77906 Human PRO polynucl
20	1608	100.0	1570	25	ABX79502 Human secreted/tr
21	1608	100.0	1570	25	ABX64141 cDNA encoding huma
22	1608	100.0	1570	25	ABX17105 Human PRO polynucl
23	1595	99.2	1476	20	AAZ16295 Human kallikrein e
24	1513	94.1	1439	20	AAZ06260 Human secreted pro
25	1403	87.3	1504	24	ABK30231 Human G-protein-co
26	1002	62.3	11570	21	AAZ95905 Human KLK-L2 gene.
27	734	45.6	455	20	AAZ57986 Human BS247 specif
28	725	45.1	735	21	AAZ79469 cDNA sequence of h
29	725	45.1	735	24	ABK29013 Human breast tumou
30	725	45.1	738	20	ABK24240 DNA encoding human
31	688	42.8	673	22	AAF44901 Human breast cance
32	671	41.7	765	22	AAZ63914 Human prostate cDN
33	671	41.7	765	22	AAH93821 Human prostate-spe
34	671	41.7	765	22	AAH85135 Human prostate-spe
35	671	41.7	765	22	ABU95285 Human P703P putati
36	671	41.7	765	25	ACA59722 Prostate cancer th
37	671	41.7	765	25	ABQ83343 Human KLK4 encodin
38	654.5	40.7	1052	21	AAZ87798 Activation constru
39	654.5	40.7	1052	22	AAZ55270 Nucleotide sequenc
40	652.5	40.6	1146	20	AAV84589 Human secreted pro
41	652.5	40.6	1146	22	ABA83430 Human secreted pro
42	652.5	40.6	1166	22	AAZ14841 Human PS133 consen
43	650.5	40.5	833	19	AAV42925 DNA encoding a hum
44	650.5	40.5	1106	20	AAZ22638 CASB12 nucleotide
45	650.5	40.5	1158	20	AAZ22639 CASB12 derived fro

ALIGNMENTS

RESULT 1  
AAZ23318  
ID AAZ23318 standard; DNA; 882 BP.  
AC AAZ23318;  
XX  
XX 03-DEC-1999 (first entry)  
XX  
XX Human PDSP-1 DNA coding region.  
XX

PDSP-1; Tango 114; prostate-derived serine protease; fusion protein;  
treatment; proliferative disorder; prostate cancer; screening;  
predictive medicine; growth factor; growth factor biosynthesis;  
cellular proliferation; growth factor binding protein; metastasis;  
cellular differentiation; prostate development; detection; forensic;  
serine protease; human; ds.  
XX

OS Homo sapiens.  
 PN WO9946391-A2.  
 XX  
 PD 16-SEP-1999.  
 XX  
 PF 12-MAR-1999; 99WO-US05416.  
 XX  
 PR 12-MAR-1998; 98US-0041400.  
 XX  
 PA (MILL-) MILLENNIUM BIOTHERAPEUTICS INC.  
 PI Holtzman DA;  
 XX  
 DR WPI; 1999-561678/47.  
 DR P-PSDB; AAY30524.  
 XX  
 PT Novel prostate derived serine protease polynucleotides and polypeptides  
 PT used to modulate cellular processes  
 XX  
 PS Claim 1a; Fig 1A; 94pp; English.  
 XX  
 CC This invention describes a novel human prostate-derived serine protease  
 CC (PDSP) polynucleotide (also known as Tango 114). The prostate derived  
 CC serine protease (PDSP) polypeptide is used to modulate a variety of  
 CC cellular processes. It can be used to produce fusion proteins. PDSP is used  
 CC to treat proliferative disorders, e.g. prostate cancer. The protein may  
 CC also be used to produce antibodies, and to identify antagonists and  
 CC agonists. The PDSP polynucleotides, polypeptides, homologs and antibodies  
 CC can be used in screening assays; predictive medicine; and methods of  
 CC treatment. PDSP cleaves growth factors, and can be used for the  
 CC modulation of growth factor biosynthesis; generation of active peptides;  
 CC regulation of cellular proliferation; degradation of growth factor  
 CC binding proteins; regulation of cellular differentiation; regulation of  
 CC metastasis; and regulation of prostate development. The PDSP  
 CC polynucleotides can be used to express the protein; to detect PDSP mRNA;  
 CC to detect genetic alterations in the PDSP gene; in forensic biology; and  
 CC as a source of primers and probes. As serine protease have important  
 CC roles in cellular processes, there exists a need for identifying novel  
 CC serine protease such as the prostate derived serine protease (PDSP) of  
 CC the invention. This sequence encodes the human PDSP-1 described in the  
 CC invention.

XX SQ Sequence 882 BP; 193 A; 272 C; 237 G; 180 T; 0 other;

Alignment Scores:  
 Pred. No.: 9,77e-138 Length: 882  
 Score: 1608.00 Matches: 293  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 20 Gaps: 0

US-09-936-271B-14 (1-293) x AAZ23318 (1-882)

QY 1 MetaAlaThrAlaArgProProTrrpMetTrrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
 DB 1 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTGCTGTGATCAGACGCTTGTCT 60  
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
 DB 61 CTGGGGGTTCACAGAGCATGTTCTCGCAACAAATGATGTTCTCTGTGACCCCTCTAAC 120  
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
 DB 121 ACCGTGCCCTCTGGAGCAACAGACCTGGGAGCTGGGGCCGGGAGAGCCCGGTGCG 180  
 QY 61 AspAspSerSerArgIleAlaAsnGlySerAspCysAspMetHisThrGlnProTrrp 80  
 DB 181 GATGACAGCAGCAGCCGCATCATCAATGATCCGACTGCGATATGCGACACCCACCGCTGG 240  
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTrrpCysGlyAlaValLeuValHisPro 100  
 |||||||

241 CAGCCCGCGCTGTTGCTAAGGCCCAACAGCTCTACTCGGGGGCGGTGTGGTGCATCCA 300  
 QY 101 GlnTrrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120  
 DB 301 CAGTGGCTGCTACGGCCGCCCTGCTGAGGAAGAAGTTTTCAGAGTCCGCTCTCGGCAC 360  
 QY 121 TyrSerLeuSerProValTrrpGluSerGlyGlnGlnMetPheGlnGlyValLysSerile 140  
 DB 361 TACTCCCTGTCAACAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGGTCAAACTCATC 420  
 QY 141 ProHisProGlyTrrpSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
 DB 421 CCCACCCCTGGCTACTCCACCCCTGGCAGCTTAAAGATGTGAGATGCTCAATGCTCTCT 480  
 QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerHisSerHisCysProSer 180  
 DB 481 AGAAGAATTCGTCCCACTAAAGATGTGAGATGCTCAATGCTCTCTCTCTCTCTCTCTCT 540  
 QY 181 AlaGlyThrLysCysLeuValSerGlyTrrpGlyThrThrLysSerProGlnValHisPhe 200  
 DB 541 GCTGGACAAAGTCTTGGTGTCTGCTGGGGGACACCAAGAGACCCCAAGTGCACCTTC 600  
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
 DB 601 CCTAAGGTCCTCCAGTGTCTTGAATATCAGCGTGTAAAGTCAGAAAGGTGCGAGATGCT 660  
 QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
 DB 661 TACCCGAGACAGATAGATGACCCATGTTCTGCGCCCGGTGACAAGCAGGTAGAGACTCC 720  
 QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
 DB 721 TGCCAGGGTGATTCCTGGGGGCCCTGTGTGCTGCAATGGCTCCCTGCAGGAGCTGTGTC 780  
 QY 261 TrrpGlyAspTrrpProCysAlaArgProAsnArgProGlyValTrrpThrAsnLeuCysLys 280  
 DB 781 TGGGAGATTACCTTGTGCGCCGCCCAACAGACGGGTGTCTACACGAACCTCTGCAAG 840  
 QY 281 PheThrLysTrrpIleGlnGluThrIleGlnAlaAsnSer 293  
 DB 841 TTCACCAAGTGGATCCAGGAACCAATCCAGGCCCAACTCC 879

RESULT 2  
 AAX57989  
 ID AAX57989 standard; DNA; 1302 BP.  
 XX  
 AC AAX57989;  
 XX  
 DT 19-JUL-1999 (first entry)  
 XX  
 DE Human BS247 specific polynucleotide #7.  
 XX  
 KW BS247; detection; diagnosis; breast cancer; atypical hyperplasia;  
 KW fibroadenoma; cystic breast disease; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9922027-A1.  
 XX  
 PD 06-MAY-1999.  
 XX  
 PF 28-OCT-1998; 98WO-US22906.  
 XX  
 PR 28-OCT-1997; 97US-0968838.  
 PR 28-OCT-1997; 97US-0063431.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 XX  
 PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;  
 PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;  
 PI Russell JC, Stroupe SD, Yu H;  
 XX  
 DR WPI; 1999-312977/26.

XX  
PT Breast tissue derived cDNA contig and consensus polypeptide sequence  
XX  
PS Claim 1; Page 105; 112pp; English.  
XX  
CC This sequence is a BS247 specific polynucleotide.  
CC The invention relates to a method of detecting the presence of a target  
CC BS247 polynucleotide, especially mRNA, in a test sample. BS247  
CC polynucleotides are derived from breast tissue. The polynucleotides,  
CC polypeptides or antibodies are useful for providing information leading  
CC to the detection, diagnosis, staging, monitoring, prognosis, in vivo  
CC imaging, prevention or treatment, determining predisposition to, diseases  
CC and conditions of the breast, such as breast cancer, atypical  
CC hyperplasia, fibroadenoma and cystic breast disease. Drug treatment or  
CC gene therapy for breast cancer, can be based on these identified gene  
CC sequences and the efficacy of any particular therapy can be monitored.  
CC The BS247-derived reagents are advantageous for detection of breast  
CC cancer due to their specificity. The reagents also provide an  
CC alternative, non-surgical diagnostic method capable of detecting early  
CC stage breast disease, such as cancer.  
XX  
SQ Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 other;

Alignment Scores:  
Pred. No.: 1,62e-137 Length: 1302  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0

US-09-936-271B-14 (1-293) x AAX57989 (1-1302)

QY 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
DB 88 ATGGCTACAGCAAGACCCCTGGATGGTGGTCTCTGGCTCTGATCATCACCCCTGCTT 147  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
DB 148 CTGGGGGTTCACAGAGCATGTTCTGCCAACATGATGTTCTCTGACCACTCTTAAC 207  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
DB 208 ACCGTGCCCTCTGGAGCAACAGGACCTGGAGCTGGGGCGGGGAGACCCCGGTG 267  
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80  
DB 268 GATGACAGCAGCGCGCATCATCATGATGATGATGATGATGATGATGATGATGATG 327  
QY 81 GlnAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
DB 328 CAGGCGCGCTGTTGCTAAGGCCCAACAGCTCTACTGCGGGCGGTGTTGTCATCCA 387  
QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120  
DB 388 CAGTGGCTGCTCAGCGCGCCCACTGCAGAGAAAGTTTTCAGAGTCCGTCTCGGCCAC 447  
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
DB 448 TACTCCCTGTCACCAAGTTTATGAATCTGGGACAGATGTTCCAGGGGTCAAAATCCATC 507  
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
DB 508 CCCCACCTGGTACTCCACCTGCCACTCTACGACCTCATGCTCATCAAACTGAAC 567  
QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
DB 568 AGAAGAATTCGTCCCACTAAAGATGTCAGACCATCAAGCTCTCTCTATTGTCCTCT 627  
QY 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200  
DB 628 GCTGGGACAAAGTGTGCTGGTGGTGGGGGACAAACCAAGAGCCCCCAAGTGCATCTC 687

QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
DB 688 CCTAAGGCTCTCCAGTGTGAATATCAGCGTGAAGTCAGAAAAGGTGCCGAGATGCT 747  
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
DB 748 TACCCGAGACAGATAGATGACACCATGTTCTCGCGCGGTGACAAAAGCAGGTAGACTCC 807  
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
DB 808 TGCCAGGTGATTCCTGGGGGGCTGTGGTCTGCATGGCTCCTGCAGGGACTCGTGTC 867  
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
DB 868 TGGGAGAGATTACCTTGTGCGCGGCCAACAGACGGGTGTCTACACGAACCTCTGCAAG 927  
QY 281 PheThrLysTrpIleGlnGlnThrIleGlnAlaAsnSer 293  
DB 928 TTCACCAAGTGGATCCAGGAACCATCCAGGCCAACTCC 966  
RESULT 3  
AAX57990  
ID AAX57990 standard; DNA; 1302 BP.  
XX  
AC AAX57990;  
XX  
DT 19-JUL-1999 (first entry)  
XX  
DE Human BS247 specific polynucleotide #8.  
XX  
KW BS247; detection; diagnosis; breast cancer; atypical hyperplasia;  
KW fibroadenoma; cystic breast disease; gene therapy; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO922027-A1.  
XX  
PD 06-MAY-1999.  
XX  
PF 28-OCT-1998; 98WO-US22906.  
XX  
PR 28-OCT-1997; 97US-0968838.  
PR 28-OCT-1997; 97US-0063431.  
XX  
PA (ABBO ) ABBOTT LAB.  
XX  
PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN,  
PI Gordon J, Granados EN, Hodges SC, Klass MR, Kratochvil JD;  
PI Russell JC, Stroupe SD, Yu H;  
XX  
DR WPI; 1999-312977/26.  
XX  
PT Breast tissue derived cDNA contig and consensus polypeptide sequence  
XX  
PS Claim 1; Page 106; 112pp; English.  
XX  
CC This sequence is a BS247 specific polynucleotide.  
CC The invention relates to a method of detecting the presence of a target  
CC BS247 polynucleotide, especially mRNA, in a test sample. BS247  
CC polynucleotides are derived from breast tissue. The polynucleotides,  
CC polypeptides or antibodies are useful for providing information leading  
CC to the detection, diagnosis, staging, monitoring, prognosis, in vivo  
CC imaging, prevention or treatment, determining predisposition to, diseases  
CC and conditions of the breast, such as breast cancer, atypical  
CC hyperplasia, fibroadenoma and cystic breast disease. Drug treatment or  
CC gene therapy for breast cancer, can be based on these identified gene  
CC sequences and the efficacy of any particular therapy can be monitored.  
CC The BS247-derived reagents are advantageous for detection of breast  
CC cancer due to their specificity. The reagents also provide an  
CC alternative, non-surgical diagnostic method capable of detecting early  
CC stage breast disease, such as cancer.  
XX  
SQ Sequence 1302 BP; 281 A; 413 C; 338 G; 270 T; 0 other;

## Alignment Scores:

Pred. No.: 1,62e-137 Length: 1302  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0

US-09-936-271B-14 (1-293) x AAZ57990 (1-1302)

Qy 1 MetAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuLeuThrAlaLeuLeu 20  
Db 88 ATGGCTACAGCAAGACCCCTGGATGGTGGTCTCTGTCTGTATCAGACGCTTGCTT 147  
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
Db 148 CTGGGGTTCACAGCATGTTCTCGCCCAATGATGTTCTCTGTACACCCCTCTAAC 207  
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60  
Db 208 ACCGTGCTCTGGGACCAACAGGACCTGGAGCTGGGGCCGGGAAGACGCCGGTGC 267  
Qy 61 AspAspSerSerArgIleLeuAenGlySerAspCysAspMetHisThrGlnProTrp 80  
Db 268 GATGACAGCAGCAGCCCATCATGATGATGATGATGATGATGATGATGATGATGATG 327  
Qy 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
Db 328 CAGGCCCGCTGTGCTTAAGGCCCAACACCTCTACTCGGGGGGTGTGTGTGATCCA 387  
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120  
Db 388 CAGTGGCTGCTACGGCGCCCACTGCAGGAAGAAGTTTTCAGAGTCCCTCGGCCAC 447  
Qy 121 TyrSerLeuSerProValTyrGluSerGlyClnGlnMetPheGlnGlyValLysSerIle 140  
Db 448 TACTCCCTGTACACAGTTTATGATTCGGCAGCAGATGTTCCAGGGGGTCAATCCATC 507  
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160  
Db 508 CCCACCTGGCTACTCCACCTGGCCACTCTACAGCACCTCATGCTCATCAACTGAAC 567  
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
Db 568 AGAAGAATTCGTCCCACTAAAGATGTGATGATGATGATGATGATGATGATGATGATG 627  
Qy 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
Db 628 GCTGGGACAAGTCTTGGTCTGTGGCTGGGGACAACCAAGAGCCCAAGTGCACTTC 687  
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
Db 688 CCTAAGTCTCCAGTCTTGAATATCAGCGTGTAGTCAGAAAGAGTGGAGGATGCT 747  
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
Db 748 TACCCGAGACAGATAGATGACACCATGTTCTGCGCCGGTGACAAAGCAGGTAGACTCC 807  
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
Db 808 TCCAGGGTGATTCGGGGGCGCTTGTGCTGCAATGGCTCCCTGCGAGGACTCGTGTC 867  
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
Db 868 TGGGAGATTACCTTGTGCGCCGCCCAACAGACCGGGTGTCTACAGCAACCTCTGCAAG 927  
Qy 281 PheThrLysTrpIleGlnGlnThrIleGlnAlaAsnSer 293  
Db 928 TTCACCAAGTGGATTCAGGAACCATCCAGGCCCAACTCC 966

RESULT 4  
AAZ06259

AAZ06259 standard; DNA; 1381 BP.

AAZ06259;

30-SEP-1999 (first entry)

Human secreted protein gene No. 27.

Human; secreted protein; fusion protein; gene therapy; protein therapy;  
diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.

Homo sapiens.

WO9935158-A1.

15-JUL-1999.

06-JAN-1999; 99WO-US00108.

07-JAN-1998; 98US-0070704.

07-JAN-1998; 98US-0070657.

07-JAN-1998; 98US-0070658.

07-JAN-1998; 98US-0070692.

(HUMA-) HUMAN GENOME SCI INC.

Brewer LA, Duan RD, Ebner R, Lafleur DW, Ni J;

Olson HS, Rosen CA, Ruben SM, Soppet DR;

WPI; 1999-444190/37.

P-PSDB; AAY38426.

New isolated human genes and the secreted polypeptides they encode

Claim 1; Page 173-174; 227pp; English.

This sequence represents a nucleic acid molecule which encodes a  
secreted human protein. The gene number is given in the descriptor line.  
The gene can be used to generate fusion proteins by linking to the gene  
to a human immunoglobulin Fc portion (e.g. AAZ06210) for increasing the  
stability of the fused protein as compared to the human protein only.  
The invention relates to 36 novel genes and their fragments (nucleic  
acid sequences: AAZ06219-206263; amino acid sequences AAY3886-Y38498)  
which are useful for preventing, treating or ameliorating medical  
conditions e.g. by protein or gene therapy. Also, pathological  
conditions can be diagnosed by determining the amount of the new  
polypeptides in a sample or by determining the presence of mutations in  
the new polynucleotides. Specific uses are described for each of the 36  
polynucleotides, based on which tissues they are most highly expressed in  
(see AAZ06219 for described uses).

Sequence 1381 BP; 329 A; 425 C; 353 G; 274 T; 0 other;

## Alignment Scores:

Pred. No.: 1-75e-137 Length: 1381  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0

US-09-936-271B-14 (1-293) x AAZ06259 (1-1381)

Qy 1 MetAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuLeuThrAlaLeuLeu 20  
Db 129 ATGGCTACAGCAAGACCCCTGGATGGTGGTCTCTGTCTGTATCAGACGCTTGCTT 188

QY	21	LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn	40
Db	189	CTGGGGTACAGAGCATGTCGCCAACAAATGATGTTCTGTGACCACTTAAC	248
QY	41	ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaAspAlaArgSer	60
Db	249	ACCGTGCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCCGGGAACGCCCGGTG	308
QY	61	AspAspSerSerArgIleLeuAsnGlySerAspCysAspMetHisThrGlnProTtp	80
Db	309	GATGACAGCAGCAGCGCATCATCATGATCGGATGCGATGATGACACCCAGCGTGG	368
QY	81	GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro	100
Db	369	CAGCGCGCGTGTGCTAAGGCCCAACACAGCTCTACTGCGGGCGGTGTTGGTCATCCA	428
QY	101	GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis	120
Db	429	CAGTGGCTGCTCAGCGCGCGCCACTGCAGGAAGAAAGTTTCAGAGTCCGCTCTCGGCCAC	488
QY	121	TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle	140
Db	489	TACTCCCTGTCACCACTTTATGATCTGGCAGCAGATGTTCCAGGGGTCAATCCATC	548
QY	141	ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn	160
Db	549	CCCCACCCCTGGCTACTCCACCCCTGACCTCTAACGACCTCATGCTCAACACTGAAC	608
QY	161	ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer	180
Db	609	AGAAGAATTGCTCCCACTAAGATGTCAGACCCATCAAGCTCTCTCTCTCTCTCTCT	668
QY	181	AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe	200
Db	669	GCTGGGCAAAAGTCTGTGTGCTGGTGGGACCAACCAAGAGCCCCCAAGTGCATTC	728
QY	201	ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla	220
Db	729	CCTAAGTCTCTCCAGTCTTGAATATCAGCGTCTAAGTCAGAAAAAGGTGCGAGGATGCT	788
QY	221	TyrProArgGlnIleAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer	240
Db	789	TACCCGAGACAGATAGATACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCC	848
QY	241	CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer	260
Db	849	TGCCAGGGTGATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGAGGAGACTCGTGTC	908
QY	261	TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys	280
Db	909	TGGGGAGATTACCTTGTGCGCGGCCAACAGACCGGGTGTCTACACGAACCTCTGCAAG	968
QY	281	PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer	293
Db	969	TTCAACCAAGTGGATCCAGGAACCACTCCAGGCCAATCC	1007
RESULT 5			
AA60578			
ID	AA60578	standard; DNA; 1499 BP.	
XX	AC	AA60578;	
XX	XX		
DT	02-AUG-1999	(first entry)	
XX	XX	Human keratinocyte derived protease (KDP) encoding DNA.	
XX	XX	Keratinocyte derived protease; KDP; protease; keratinocyte; human;	
KW	KW	skin care product; skin flaking; dandruff; laundry detergent composition;	
XX	XX	cleaning composition; dishwashing product; ss.	
OS	OS	Homo sapiens.	
XX	XX	Key	
FH		Location/Qualifiers	

FT	5' UTR	1..290	
FT		/*tag- a	
FT	CDS	291..1172	
FT		/*tag- b	"keratinocyte derived protease"
FT		/product-	291..488
FT	sig_peptide	/*tag- c	489..1169
FT	mat_peptide	/*tag- d	1173..1499
FT	3' UTR	/*tag- e	
XX		WO9918219-A1.	
XX		15-APR-1999.	
XX		03-OCT-1997;	97WO-US17864.
XX		03-OCT-1997;	97WO-US17864.
XX	(PROC )	PROCTER & GAMBLE CO.	
XX		Kitado H, Yoshikawa A, Zaiki T;	
XX		WPI; 1999-287737/24.	
XX		P-PSDB; AAV16777.	
XX		Novel human protease useful for treating or preventing skin flaking	
XX	Claim 3;	Page 32-34; 43pp; English.	
XX		This DNA encodes a human keratinocyte derived protease (KDP) polypeptide.	
XX		The protease is substantially similar to a sequence encoded by the insert	
XX		in plasmid FERM BP-6129. The KDP polypeptide, from human keratinocytes,	
XX		is specifically used in skin care products, particularly to treat or	
XX		prevent skin flaking, e.g. dandruff, and in laundry detergent	
XX		compositions. More generally the KDP polypeptide can be used in any	
XX		cleaning composition, e.g. hard surface or dishwashing products.	
XX	Sequence	1499 BP; 310 A; 469 C; 398 G; 322 T; 0 other;	
XX	Alignment Scores:		
XX	Pred. No.:	1.95e-137	Length: 1499
XX	Score:	1608.00	Matches: 293
XX	Percent Similarity:	100.00%	Conservative: 0
XX	Best Local Similarity:	100.00%	Mismatches: 0
XX	Query Match:	100.00%	Indels: 0
XX	DB:	20	Gaps: 0
XX		US-09-936-271B-14 (1-293) x AAX60578 (1-1499)	
QY	1	MetAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu	20
Db	291	ATGGCTACAGCAAGACCCCTGGATGTGGTGCTGTGCTGTGATCAGCGCTTGCTT	350
QY	21	LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn	40
Db	351	CTGGGGTCCACAGAGCATGTTCTCGCCAAACATGATGTTCTGTGACCACTTAAC	410
QY	41	ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer	60
Db	411	ACCGTGCCTCTGGGAGCAACAGGACCTGGGAGCTGGGCGCGGGGAAGACCGCGTGC	470
QY	61	AspAspSerSerArgIleLeuAsnGlySerAspCysAspMetHisThrGlnProTrp	80
Db	471	GATGACAGCAGCAGCGCATCATCAATGATCGGATGCGATGATGACACCCAGCGTGG	530
QY	81	GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro	100
Db	531	CAGCGCGCGCTGTGCTAAGGCCCAACAGCTCTACTGCGGGCGGTGTTGGTGCATCCA	590
QY	101	GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis	120

Db 591 CAGTGGCTGCTCAGCGCGCCGCTGCGAGGAAGAAAGTTTTCAGAGTCCCTCTCGGGCCAC 650  
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
Db 651 TACTCCCTGTGTCACCACTTATGATCTGGCAGCAGATGTTCCAGGGGGTCAATCCATC 710  
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160  
Db 711 CCCACCGCTGGCTACTCCCGCTGGCCACTTAACGACCTCATGCTCATCAAACTGAAC 770  
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
Db 771 AGAAGAATCGTCCCACTAAAGATGTCAGACCCATCAACGCTCTCTCATTTGCTCTCT 830  
Qy 181 AlaGlyThrLysCysLeuValSerGlyTyrPheGlyThrThrLysSerProGlnValHisPhe 200  
Db 831 GCTGGGACAAAGTCTGGTGTCTGGCTGGGGACACCAAGAGCCCGCCCAAGTGCATTC 890  
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
Db 891 CCTAAGGTCTCCAGTCTTGAATATCAGCGGTGCTAAGTCAAGAAAGTGGCAGGATGCT 950  
Qy 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
Db 951 TACCCGAGACAGATGATGACACCATGTTCTGCCCGGTGACAAAGCGGTAGAGATCC 1010  
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
Db 1011 TGCCAGGTGATTTCTGGGGGCCCTGTGCTGCAATGGCTCCCTGCGAGGACTCGTGTC 1070  
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
Db 1071 TGGGGAGATTACCCCTTGTGCCCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAG 1130  
Qy 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
Db 1131 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1169  
RESULT 6  
AAZ06245  
ID AAZ06245 standard; DNA; 1516 BP.  
XX  
AC AAZ06245;  
XX  
DT 30-SEP-1999 (first entry)  
XX  
DE Human secreted protein gene No. 27.  
XX  
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;  
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;  
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;  
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;  
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;  
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;  
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;  
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.  
XX  
OS Homo sapiens.  
XX  
PN W09J35158-A1.  
XX  
PD 15-JUL-1999.  
XX  
PF 06-JAN-1999; 99WO-US00108.  
XX  
PR 07-JAN-1998; 98US-0070704.  
PR 07-JAN-1998; 98US-0070657.  
PR 07-JAN-1998; 98US-0070658.  
PR 07-JAN-1998; 98US-0070692.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Brewer LA, Duan RD, Ebner R, Lafleur DW, Ni J;

PI Olsen HS, Rosen CA, Ruben SM, Soppet DR;  
XX  
DR WPI: 1999-444190/37.  
DR P-PSDB; AAY38412.  
XX  
PT New isolated human genes and the secreted polypeptides they encode  
XX  
PS Claim 1; Page 163-164; 227pp; English.  
XX  
CC This sequence represents a nucleic acid molecule which encodes a  
CC secreted human protein. The gene number is given in the descriptor line.  
CC The gene can be used to generate fusion proteins by linking to the gene  
CC to a human immunoglobulin Fc portion (e.g. AAZ06210) for increasing the  
CC stability of the fused protein as compared to the human proteins only.  
CC The invention relates to 36 novel genes and their fragments (nucleic  
CC acid sequences: AAZ06219-206263; amino acid sequences AAY38386-Y38498)  
CC which are useful for preventing, treating or ameliorating medical  
CC conditions e.g. by protein or gene therapy. Also, pathological  
CC conditions can be diagnosed by determining the amount of the new  
CC polypeptides in a sample or by determining the presence of mutations in  
CC the new polynucleotides. Specific uses are described for each of the 36  
CC polynucleotides, based on which tissues they are most highly expressed in  
CC (see AAZ06219 for described uses).  
XX  
SQ Sequence 1516 BP; 338 A; 472 C; 381 G; 324 T; 1 other;  
  
Alignment Scores:  
Pred. No.: 1,98e-137 Length: 1516  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0  
  
US-09-936-271B-14 (1-293) x AAZ06245 (1-1516)  
Qy 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
Db 254 ATGGCTACAGCAAGACCCCTGGATGGTGGTCTCTGTCTGTATGCACAGCTTGTCTT 313  
Qy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
Db 314 CTGGGGTTCACAGCATGTCTCGCCCAACAATGATGTTTCTCTGTGACCACTCTTAC 373  
Qy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60  
Db 374 ACCGTGCCCTCTGGGAGCAACAGGACCTGGAGCTGGGGCGGGGAGACGCCGGTCG 433  
Qy 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMethHisThrGlnProTrp 80  
Db 434 GATGACAGCAGCAGCCCATCATCAATGGATCGGACTATGCACACCCAGCGTGG 493  
Qy 81 GlnAlaAlaLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
Db 494 CAGCCCGCGTGTGCTTAAGGCCCAACCACTCTACTGCGGGCGGTGTGGTGCATCCA 553  
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120  
Db 554 CAGTGGCTGCTCAGCGCCGCCACTCGCAGGAAGAAGTTCAGAGTCCCTCTCGGGCAC 613  
Qy 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
Db 614 TACTCCCTGTGACAGTTTATGATCTGGGCAGCAGATGTTCCAGGGGGTCAATCCATC 673  
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160  
Db 674 CCCACCGCTGGCTACTCCCGCTGGCCACTCTTAACGACCTCATGCTCATCAAACTGAAC 733  
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
Db 734 AGAAGAATTCGTCCTCCACTAAAGATGTCAGACCCATCAACGCTCTCTCTCTATTGTCCTCT 793  
Qy 181 AlaGlyThrLysCysLeuValSerGlyTyrPheGlyThrThrLysSerProGlnValHisPhe 200

|||||  
794 GCTGGCAAAAGTGTCTGGTCTGGCTGGGGCAACCAAGAGCCCCCAAGTGCACCTTC 853  
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnIysArgCysGluAspAla 220  
Db 854 CTTAAGTCTCTCAGTGTCTGAATATCAGCGTGTAAAGTCAGAAAGGTGCGAGATGCT 913  
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
Db 914 TACCCGAGACAGATAGATCACACCATGTTCTGCGCGGTGACAAAGCAGGTAGACTCC 973  
QY 241 CysGlnGlyAspSerGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
Db 974 TGCCAGGGTGATTCTGGGGGGCTGTGTGCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 1033  
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
Db 1034 TGGGGAGATTACCTTGTCCCGGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAG 1093  
QY 281 PheThrLysTrpIleGlnThrIleGlnAlaAsnSer 293  
Db 1094 TTCACCAAGTGGATCCAGGAACCATCCAGGCCAATCC 1132

RESULT 7

AAZ23317

ID AAZ23317 standard; DNA; 1539 BP.

XX AC AAZ23317;

XX DT 03-DEC-1999 (first entry)

XX DE Human PDSP-1 DNA.

XX KW PDSP-1; Tango 114; prostate-derived serine protease; fusion protein; treatment; proliferative disorder; prostate cancer; screening;  
XX KW predictive medicine; growth factor; growth factor biosynthesis;  
XX KW cellular proliferation; growth factor binding protein; metastasis;  
XX KW cellular differentiation; prostate development; detection; forensic;  
XX KW serine protease; human; ds.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers  
XX CDS 244..1125  
XX FT /\*tag= a  
XX FT /product= "PDSP-1"

XX PN W09946391-A2.

XX PD 16-SEP-1999.

XX PF 12-MAR-1999; 99WO-0505416.

XX PR 12-MAR-1998; 98US-0041400.

XX PA (MILL-) MILLENNIUM BIOTHEAPEUTICS INC.

XX PI Holtzman DA;

XX DR WPI; 1999-561678/47.

XX PS P-PSDB; AAY30524.

XX PT Novel prostate derived serine protease polynucleotides and polypeptides used to modulate cellular processes -

XX PS Claim 1a; Fig 1A; 94pp; English.

XX CC This invention describes a novel human prostate-derived serine protease (PDSP) polynucleotide (also known as Tango 114). The prostate derived serine protease (PDSP) polypeptide is used to modulate a variety of cellular processes. It can be used to produce fusion proteins. PDSP is used to treat proliferative disorders, e.g. prostate cancer. The protein may also be used to produce antibodies, and to identify antagonists and

CC agonists. The PDSP polynucleotides, polypeptides, homologs and antibodies can be used in screening assays; predictive medicine; and methods of treatment. PDSP cleaves growth factors, and can be used for the modulation of growth factor biosynthesis; generation of active peptides; regulation of cellular proliferation; degradation of growth factor binding proteins; regulation of cellular differentiation; regulation of metastasis; and regulation of prostate development. The PDSP polynucleotides can be used to express the protein; to detect PDSP mRNA; to detect genetic alterations in the PDSP gene; in forensic biology; and as a source of primers and probes. As serine protease have important roles in cellular processes, there exists a need for identifying novel serine protease such as the prostate derived serine protease (PDSP) of the invention. This sequence encodes the human PDSP-1 described in the invention.

XX SQ Sequence 1539 BP; 340 A; 470 C; 413 G; 315 T; 1 other;

Alignment Scores:

Pred. No.: 2,02e-137 Length: 1539  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 20 Gaps: 0

US-09-936-271B-14 (1-293) x AAZ23317 (1-1539)

QY 1 MetAlaThrAlaArgProThrPheMetTrpValLeuCysAlaLeuIleThrAlaLeu 20  
Db 244 ATGGCTACAGCAAGACCCCTGGATGTGGGTGCTGTGCTGTATCATCAGCCTTGCTT 303  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
Db 304 CTGGGGGTACAGAGCATGTTCTCGCCACATGATGTTTCTGTGACACCCCTCTAAC 363  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
Db 364 ACCGTGCCCTCTGGGAGCAACACGAGACCTGGGAGCTGGGGCGGGGAGAGCCCGGTG 423  
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMethIsthrGlnProTrp 80  
Db 424 GATGACAGCAGCGCGCATCATCATGATCGGCTGCGATGATGACACCCAGCGGTGG 483  
QY 81 GlnAlaAlaLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
Db 484 CAGGCGCGCTGCTGCTAAGGCCCAACACAGCTCTACTGCGGGCGGTGTTGGTCATCCA 543  
QY 101 GlnTrpLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120  
Db 544 CAGTGGCTGCTCAGCGCGCCCACTGCAGAGAAAGTTTTCAGAGTCCGCTCTCGGCCAC 603  
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
Db 604 TACTCCCTGTCCACAGTTTATGAATCTGGGAGCAGATGTTCCAGGGGTCAATCATTC 663  
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
Db 664 CCCCACCTGGCTACTCCACCCCTGACACTCTAAGGACCTCATGCTCATCAACTGAAC 723  
QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
Db 724 AGAAGAATTCGTCCTCCACTAAAGATGTCAGACCATCAACGTCTCTCTCATTTGCTCCT 783  
QY 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
Db 784 GCTGGACAAAGTGTGTGTGTGCTGGGGACCAACCAAGAGCCCCCAAGTGCATTC 843  
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
Db 844 CTTAAGTCTCTCAGTGTCTGAATATCAGCGTGTAAAGTCAGAAAAGGTGCGAGGATGCT 903  
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240

```
Db 904 TACCCGACAGATAGATGACACACATGTTCTGCGCCGGTGACAAAGCAGGTAGAGACTCC 963
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db 964 TGCAGGGTGATTCGGGGGCGCTGTGCTGCAATGCTCCCTGCAGGACTCGTGCTCC 1023
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 1024 TGGGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACGAACCTCTGCAAG 1083
Qy 281 PheThrIysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db 1084 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1122

RESULT 8
AAZ65070
ID AAZ65070 standard; cDNA; 1570 BP.
XX
AC AAZ65070;
XX
DT 05-APR-2000 (first entry)
DE Membrane-bound protein PRO1132 encoding cDNA.
KW Membrane-bound polypeptide; PRO polypeptide; LDL receptor; TIE ligand;
KW pharmaceutical; receptor immunoadhesin; gene mapping; ss.
OS Homo sapiens.
PN WO9963088-A2.
XX
PD 09-DEC-1999.
XX
PF 02-JUN-1999; 99WO-US12252.
XX
PR 02-JUN-1998; 98US-0087607.
PR 02-JUN-1998; 98US-0087609.
PR 02-JUN-1998; 98US-0087759.
PR 03-JUN-1998; 98US-0087827.
PR 04-JUN-1998; 98US-0088021.
PR 04-JUN-1998; 98US-0088025.
PR 04-JUN-1998; 98US-0088028.
PR 04-JUN-1998; 98US-0088029.
PR 04-JUN-1998; 98US-0088030.
PR 04-JUN-1998; 98US-0088033.
PR 04-JUN-1998; 98US-0088326.
PR 05-JUN-1998; 98US-0088167.
PR 05-JUN-1998; 98US-0088202.
PR 05-JUN-1998; 98US-0088212.
PR 05-JUN-1998; 98US-0088217.
PR 09-JUN-1998; 98US-0088655.
PR 10-JUN-1998; 98US-0088722.
PR 10-JUN-1998; 98US-0088730.
PR 10-JUN-1998; 98US-0088734.
PR 10-JUN-1998; 98US-0088738.
PR 10-JUN-1998; 98US-0088740.
PR 10-JUN-1998; 98US-0088741.
PR 10-JUN-1998; 98US-0088742.
PR 10-JUN-1998; 98US-0088810.
PR 10-JUN-1998; 98US-0088811.
PR 10-JUN-1998; 98US-0088824.
PR 10-JUN-1998; 98US-0088825.
PR 10-JUN-1998; 98US-0088826.
PR 11-JUN-1998; 98US-0088858.
PR 11-JUN-1998; 98US-0088861.
PR 11-JUN-1998; 98US-0088863.
PR 11-JUN-1998; 98US-0088876.
PR 12-JUN-1998; 98US-0089090.
PR 12-JUN-1998; 98US-0089105.
PR 16-JUN-1998; 98US-0089440.
PR 16-JUN-1998; 98US-0089512.
PR 16-JUN-1998; 98US-0089514.
PR 17-JUN-1998; 98US-0089532.

PR 17-JUN-1998; 98US-0089538.
PR 17-JUN-1998; 98US-0089598.
PR 17-JUN-1998; 98US-0089599.
PR 17-JUN-1998; 98US-0089600.
PR 17-JUN-1998; 98US-0089653.
PR 18-JUN-1998; 98US-0089801.
PR 18-JUN-1998; 98US-0089907.
PR 18-JUN-1998; 98US-0089908.
PR 19-JUN-1998; 98US-0089947.
PR 19-JUN-1998; 98US-0089948.
PR 19-JUN-1998; 98US-0089952.
PR 22-JUN-1998; 98US-0090246.
PR 22-JUN-1998; 98US-0090252.
PR 22-JUN-1998; 98US-0090254.
PR 23-JUN-1998; 98US-0090349.
PR 23-JUN-1998; 98US-0090355.
PR 23-JUN-1998; 98US-0090355.
PR 24-JUN-1998; 98US-0090429.
PR 24-JUN-1998; 98US-0090431.
PR 24-JUN-1998; 98US-0090435.
PR 24-JUN-1998; 98US-0090444.
PR 24-JUN-1998; 98US-0090445.
PR 24-JUN-1998; 98US-0090461.
PR 24-JUN-1998; 98US-0090472.
PR 24-JUN-1998; 98US-0090535.
PR 24-JUN-1998; 98US-0090538.
PR 24-JUN-1998; 98US-0090540.
PR 24-JUN-1998; 98US-0090557.
PR 25-JUN-1998; 98US-0090676.
PR 25-JUN-1998; 98US-0090678.
PR 25-JUN-1998; 98US-0090688.
PR 25-JUN-1998; 98US-0090690.
PR 25-JUN-1998; 98US-0090691.
PR 25-JUN-1998; 98US-0090694.
PR 25-JUN-1998; 98US-0090695.
PR 25-JUN-1998; 98US-0090696.
PR 26-JUN-1998; 98US-0090862.
PR 26-JUN-1998; 98US-0090863.
PR 01-JUL-1998; 98US-0091358.
PR 01-JUL-1998; 98US-0091360.
PR 01-JUL-1998; 98US-0091544.
PR 02-JUL-1998; 98US-0091478.
PR 02-JUL-1998; 98US-0091486.
PR 02-JUL-1998; 98US-0091519.
PR 02-JUL-1998; 98US-0091626.
PR 02-JUL-1998; 98US-0091628.
PR 02-JUL-1998; 98US-0091633.
PR 02-JUL-1998; 98US-0091646.
PR 02-JUL-1998; 98US-0091673.
PR 07-JUL-1998; 98US-0091978.
PR 07-JUL-1998; 98US-0091982.
PR 09-JUL-1998; 98US-0092182.
PR 10-JUL-1998; 98US-0092472.
PR 20-JUL-1998; 98US-0093339.
PR 30-JUL-1998; 98US-0094651.
PR 04-AUG-1998; 98US-0095282.
PR 04-AUG-1998; 98US-0095285.
PR 04-AUG-1998; 98US-0095301.
PR 04-AUG-1998; 98US-0095302.
PR 04-AUG-1998; 98US-0095318.
PR 04-AUG-1998; 98US-0095321.
PR 04-AUG-1998; 98US-0095325.
PR 10-AUG-1998; 98US-0095916.
PR 10-AUG-1998; 98US-0095929.
PR 10-AUG-1998; 98US-0096012.
PR 11-AUG-1998; 98US-0096143.
PR 11-AUG-1998; 98US-0096146.
PR 12-AUG-1998; 98US-0096329.
PR 17-AUG-1998; 98US-0096757.
PR 17-AUG-1998; 98US-0096766.
PR 17-AUG-1998; 98US-0096768.
PR 17-AUG-1998; 98US-0096773.
PR 17-AUG-1998; 98US-0096791.
PR 17-AUG-1998; 98US-0096867.
```

PR 17-AUG-1998: 98US-0096891.  
 PR 17-AUG-1998: 98US-0096894.  
 PR 17-AUG-1998: 98US-0096895.  
 PR 17-AUG-1998: 98US-0096897.  
 PR 18-AUG-1998: 98US-0096949.  
 PR 18-AUG-1998: 98US-0096950.  
 PR 18-AUG-1998: 98US-0096959.  
 PR 18-AUG-1998: 98US-0096960.  
 PR 18-AUG-1998: 98US-0097022.  
 PR 19-AUG-1998: 98US-0097141.  
 PR 20-AUG-1998: 98US-0097218.  
 PR 24-AUG-1998: 98US-0097661.  
 PR 26-AUG-1998: 98US-0097951.  
 PR 26-AUG-1998: 98US-0097952.  
 PR 26-AUG-1998: 98US-0097954.  
 PR 26-AUG-1998: 98US-0097955.  
 PR 26-AUG-1998: 98US-0097971.  
 PR 26-AUG-1998: 98US-0097974.  
 PR 26-AUG-1998: 98US-0097978.  
 PR 26-AUG-1998: 98US-0097979.  
 PR 26-AUG-1998: 98US-0097986.  
 PR 26-AUG-1998: 98US-0098014.  
 PR 31-AUG-1998: 98US-0098525.  
 PR 16-SEP-1998: 98US-0100634.  
 PR 12-JAN-1999: 99US-0115565.  
 XX  
 PA (GETH ) GENENTECH INC.

PI Baker K, Chen J, Goddard A, Gurney AL, Smith V, Watanabe CK;

PI Wood WI, Yuan J;  
 XX  
 DR WPI; 2000-072883/06.  
 DR P-PSDB; AAV66726.

PT Membrane-bound proteins and related nucleotide sequences -  
 XX  
 PS Claim 2; Fig 225; 822pp; English.

XX The invention provides membrane-bound PRO polypeptides and  
 CC polynucleotides encoding them. The PRO sequences of the invention were  
 CC identified based on extracellular domain homology screening. The PRO  
 CC sequences have homology with proteins including LDL receptors, TIE  
 CC ligands and various enzymes. The membrane-bound proteins and receptor  
 CC molecules are useful as pharmaceutical and diagnostic agents. Receptor  
 CC immunoadhesins, for instance, can be used as therapeutic agents to block  
 CC receptor-ligand interactions. The membrane-bound proteins can also be  
 CC employed for screening of potential peptide or small molecule inhibitors  
 CC of the relevant receptor/ligand interaction. The PRO encoding sequences  
 CC are useful as hybridization probes, in chromosome and gene mapping and in  
 CC the generation of antisense RNA and DNA. PRO nucleic acid sequences  
 CC will also be useful for the preparation of PRO polypeptides, especially  
 CC by recombinant techniques.

XX SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

Alignment Scores:

Pred. No.: 2,07e-137 Length: 1570  
 Score: 1608.00 Matches: 293  
 Percent Similarity: 100.0% Conservative: 0  
 Best Local Similarity: 100.0% Mismatches: 0  
 Query Match: 100.0% Indels: 0  
 DB: 21 Gaps: 0

US-09-936-271B-14 (1-293) x AAZ65070 (1-1570)

QY 1 MetalAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
 DB 354 ATGGCTACAGCAAGACCCCTGGATGTGGGTGCTGTGCTGATCAGCCITGCTT 413  
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
 DB 414 CTGGGGGTACAGAGCATGTTCTGCCCAACAATGATGTTTCTGTGACCAACCCCTTAAC 473

QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
 DB ACCGTGCCCTCTGGGAGCAACACGACCTGGGAGCTGGGGCGGGGAGCAACGCCCGGTG 533  
 QY 61 AspAspSerSerArgIleLeuAsnGlySerAspCysAspMetHisThrGlnProTrp 80  
 DB GATGACAGCAGCAGCGCATCATCAATGATCGGACTCGGATGATGATGATGATGATGATG 593  
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
 DB CAGGGCGCGCTGTGCTAAGGCCCAACACGACTCTACTGGGGCGGTGTGGTGCATCCA 653  
 QY 101 GlnTrpLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120  
 DB CAGTGGCTGCTCAGGGCCGCCACTGTCAGAGAGAAGTTTTCAGAGTCCGCTCTCGGCCAC 713  
 QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
 DB TACTCCCTGTCCACAGTTTATGAATCTGGGCAGCAGATGTTCCAGGGGTCAAAATCCATC 773  
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
 DB CCCACCCCTGGCTACTCCACCTGGCCACTTACGACCTCATGCTCATCAACATGAAC 833  
 QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
 DB AGAAGNATTCGTCCTCCACTAAAGATGTCAGACCCATCAAGTCTCTCTCATTTGCTCCTCT 893  
 QY 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
 DB GCTGGGACAAAGTGTGTGTGTGGTGGGGACAAACCAAGAGGCCCAAGTGCACATTC 953  
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
 DB CTTAAGGTCTCTCCAGTGTGAATATCAGGTGCTAAGTCAGAAAAAGGTGCGAGGATGCT 1013  
 QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
 DB TACCCGAGACATAGATGACACCATGTTCTGGCGGGTGACAAACAGGATAGAGACTCC 1073  
 QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
 DB TGCCAGGGTGATTTCTGGGGGCTGTGTGTGCAATGGCTCCCTGCGAGGACTCGTGTCC 1133  
 QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
 DB TGGGGAGATTACCCCTGTGTCGGGGCCCAACAGACCGGGGTCTTACACGAACCTCTCAAG 1193  
 QY 281 PheThrLysTrpIleGlnThrIleGlnAlaAsnSer 293  
 DB TTTACCAACAGTGGATCCAGGAACCATCCAGGCCAATCC 1232  
 DB  
 RESULT 9  
 AAS21471  
 ID AAS21471 standard; cdNA; 1570 BP.  
 XX  
 AC AAS21471;  
 XX  
 DT 24-OCT-2001 (first entry)  
 XX  
 DE Human cdNA sequence encoding for PRO1132 polypeptide.  
 XX  
 KW Human secretory and transmembrane; PRO; mammalian; cancer; lung;  
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;  
 KW cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;  
 KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.  
 OS Homo sapiens.  
 XX  
 PN WO200140466-A2.  
 XX  
 PD 07-JUN-2001.

PF 01-DEC-2000; 2000WO-US32678.  
 XX 01-DEC-1999; 99WO-US28301.  
 PR 01-DEC-1999; 99WO-US28634.  
 PR 02-DEC-1999; 99WO-US28551.  
 PR 02-DEC-1999; 99WO-US28564.  
 PR 02-DEC-1999; 99WO-US28565.  
 PR 09-DEC-1999; 99US-0170262.  
 PR 16-DEC-1999; 99WO-US30095.  
 PR 20-DEC-1999; 99WO-US30911.  
 PR 20-DEC-1999; 99WO-US30999.  
 PR 30-DEC-1999; 99WO-US31243.  
 PR 06-JAN-2000; 2000WO-US00277.  
 PR 06-JAN-2000; 2000WO-US00376.  
 PR 11-FEB-2000; 2000WO-US03565.  
 PR 18-FEB-2000; 2000WO-US04341.  
 PR 18-FEB-2000; 2000WO-US04342.  
 PR 22-FEB-2000; 2000WO-US04114.  
 PR 24-FEB-2000; 2000WO-US04914.  
 PR 24-FEB-2000; 2000WO-US05004.  
 PR 01-MAR-2000; 2000WO-US05601.  
 PR 20-MAR-2000; 2000WO-US07377.  
 PR 21-MAR-2000; 2000WO-US07532.  
 PR 30-MAR-2000; 2000WO-US08439.  
 PR 17-MAY-2000; 2000WO-US13705.  
 PR 22-MAY-2000; 2000WO-US14042.  
 PR 30-MAY-2000; 2000WO-US14941.  
 PR 02-JUN-2000; 2000WO-US15264.  
 PR 10-NOV-2000; 2000WO-US30873.  
 XX

(GETH ) GENENTECH INC.

XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
 XX

DR WPI: 2001-408281/43.  
 DR P-PSDB; AAU12399.

XX Isolated, secretory and transmembrane PRO polypeptide used to detect  
 PT other PRO polypeptides, link bioactive molecules to cells expressing  
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.  
 PT lung, breast, prostate, cervical

XX Claim 3; Fig 455; 813pp; English.

XX AAS21244-AAS21518 encode for novel human secretory and transmembrane  
 CC PRO polypeptides. The PRO polypeptides are useful to detect other  
 CC PRO polypeptides, to link bioactive molecules to cells expressing  
 CC PRO polypeptides, to modulate biological activities of cells expressing  
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,  
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO  
 CC polypeptide expression in a cell sample to that in a control sample.  
 CC Some of the 275 sequences are also useful to stimulate the release of  
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the  
 CC proliferation or differentiation of chondrocytes, the proliferation or  
 CC gene expression in pericyte cells, the release of proteoglycans from  
 CC cartilage, the proliferation of inner ear utricular supporting cells or  
 CC T-lymphocytes, the release of a cytokine from peripheral blood  
 CC monocytes (PMCs), or the proliferation of endothelial cells. Some of  
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by  
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide  
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify  
 CC molecules involved in binding interactions. The polynucleotides encoding  
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,  
 CC transgenic or knock out animals and can be used in gene therapy.  
 XX

SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

Alignment Scores:

Pred. No.: 2,07e-137 Length: 1570  
 Score: 1608,00 Matches: 293  
 Percent Similarity: 100,00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 22 Gaps: 0  
 US-09-936-271B-14 (1-293) x AAS21471 (1-1570)  
 QY 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
 DB 354 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTCTGTATCAGACCTTGTCTT 413  
 QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
 DB 414 CTGGGGTCACAGACATGTTCTGCCCAACATGATGTTTCTCTGTGACCACTCTAAC 473  
 QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
 DB 474 ACCGTGCCCTCTGGGAGCAACAGGACCTGGAGCTGGGGCCGGGAAGACGCCGGTGC 533  
 QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80  
 DB 534 GATGACAGCAGCAGCCGCATCATCAATGGATCGACTGCGATATGCACACCCAGCCGTGG 593  
 QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
 DB 594 CAGCGCGCTGTGTCTTAAGCCCCAACCACTCTACTGCGGGGGGTGTGGTGCATCCA 653  
 QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis 120  
 DB 654 CAGTGGTGTCTCAGCGCGCCCACTGCAGGAAGAAAGTTTTCAGAGTCCGCTCGGCCAC 713  
 QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
 DB 714 TACTCCCTGTACACAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGGTCAATCCATC 773  
 QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
 DB 774 CCCACCTCGCTACTCCCACTTGGCCACTTAACGACCTCATGCTCATCAAACTGAAC 833  
 QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
 DB 834 AGAAGAATTCGTCCCACTAAAGATGTCCAGACCCATCAACGCTCTCTCATTTGCCCTCT 893  
 QY 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200  
 DB 894 GCTGGGACAAAGTCTTGGTGTCTGGTGGGGGACAAACCAAGACCCCAAGTGCACCTTC 953  
 QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
 DB 954 CCTAAGGTCTCTCCAGTCTTGAATATCAGCGTGTCTAAGTCAAGAAAGTGGAGGATGCT 1013  
 QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
 DB 1014 TACCCGAGACAGATAGATGACACCATGTTCTGGCGCGGTGACAAAGCAGGTAGAGACTCC 1073  
 QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
 DB 1074 TGGCAGGGTGATTTCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCGAGGACTGCTGCC 1133  
 QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
 DB 1134 TGGGGAGATTACCTTGTGCGCGCGCCCAACAGACCCGGGTGCTACACGAACCTCTGCAAG 1193  
 QY 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
 DB 1194 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1232

RESULT 10

AAF44216

ID AAF44216 standard; cDNA; 1570 BP.

XX AAF44216;

XX AAF44216;

DT 02-APR-2001 (first entry)





Db 1074 TCCAGGAGTATCTGGGGGGCCCTGTGTCTGCAATGGCTCCCTCAGGGACTCGTGCC 1133  
QY 261 TTPGLYASPTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
|||||  
Db 1134 TGGGAGATTACCTTGTGGCGGGCCCAACAGACGGGTGTCTACACGAACCTCTGCAG 1193  
QY 281 PheThrLysTyrIleGlnGluThrIleGlnAlaAsnSer 293  
|||||  
Db 1194 TTCACCAAGTGTATCCAGAAACCATCCAGGGCCAACTCC 1232  
RESULT 12  
ACA03830  
ID ACA03830 standard; cDNA; 1570 BP.  
XX AC  
ACA03830;  
XX AC  
XX DT  
XX 23-MAY-2003 (first entry)  
XX cDNA encoding human PRO polypeptide #228.  
XX Human; PRO polypeptide; secreted and transmembrane protein;  
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;  
KW differentiation; chondrocyte; tumour; genetic disorder;  
KW cytosstatic; gene; ss.  
XX Homo sapiens.  
XX US2003036180-A1.  
XX PD 20-FEB-2003.  
XX PF 09-MAY-2002; 2002US-0143114.  
XX 31-MAR-1997; 97WO-US05230.  
PR 12-JUN-1998; 98WO-US12456.  
PR 14-JUL-1998; 98WO-US14552.  
PR 28-AUG-1998; 98WO-US17888.  
PR 10-SEP-1998; 98WO-US18824.  
PR 14-SEP-1998; 98WO-US19093.  
PR 14-SEP-1998; 98WO-US19094.  
PR 14-SEP-1998; 98WO-US19177.  
PR 16-SEP-1998; 98WO-US19330.  
PR 17-SEP-1998; 98WO-US19437.  
PR 07-OCT-1998; 98WO-US21141.  
PR 29-OCT-1998; 98WO-US22991.  
PR 29-OCT-1998; 98WO-US22992.  
PR 20-NOV-1998; 98WO-US24855.  
PR 01-DEC-1998; 98WO-US25108.  
PR 05-JAN-1999; 99WO-US00106.  
PR 08-MAR-1999; 99WO-US05028.  
PR 20-APR-1999; 99WO-US08615.  
PR 14-MAY-1999; 99WO-US10733.  
PR 02-JUN-1999; 99WO-US12252.  
PR 01-SEP-1999; 99WO-US20111.  
PR 08-SEP-1999; 99WO-US20594.  
PR 13-SEP-1999; 99WO-US20944.  
PR 15-SEP-1999; 99WO-US21090.  
PR 15-SEP-1999; 99WO-US21547.  
PR 05-OCT-1999; 99WO-US23089.  
PR 29-NOV-1999; 99WO-US28214.  
PR 30-NOV-1999; 99WO-US28313.  
PR 30-NOV-1999; 99WO-US28409.  
PR 01-DEC-1999; 99WO-US28301.  
PR 01-DEC-1999; 99WO-US28634.  
PR 02-DEC-1999; 99WO-US28551.  
PR 02-DEC-1999; 99WO-US28564.  
PR 02-DEC-1999; 99WO-US28565.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 20-DEC-1999; 99WO-US30999.  
PR 22-DEC-1999; 99WO-US30720.  
PR 30-DEC-1999; 99WO-US31243.

PR 30-DEC-1999; 99WO-US31274.  
PR 05-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00277.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 18-FEB-2000; 2000WO-US04342.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 01-MAR-2000; 2000WO-US05601.  
PR 02-MAR-2000; 2000WO-US05746.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 10-MAR-2000; 2000WO-US06319.  
PR 15-MAR-2000; 2000WO-US06884.  
PR 20-MAR-2000; 2000WO-US07377.  
PR 21-MAR-2000; 2000WO-US07532.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 11-AUG-2000; 2000WO-US22031.  
PR 23-AUG-2000; 2000WO-US23522.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 10-NOV-2000; 2000WO-US30873.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 20-DEC-2000; 2000WO-US34956.  
PR 28-FEB-2001; 2001WO-US06520.  
PR 01-MAR-2001; 2001WO-US06666.  
PR 25-MAY-2001; 2001WO-US17092.  
PR 01-JUN-2001; 2001WO-US17800.  
PR 20-JUN-2001; 2001WO-US19692.  
PR 22-JUN-2001; 2001WO-US20116.  
PR 29-JUN-2001; 2001WO-US21066.  
PR 09-JUL-2001; 2001WO-US21735.  
PR 20-DEC-2000; 2000US-0747259.  
PR 28-FEB-2001; 2001US-0796498.  
PR 09-MAR-2001; 2001US-0802706.  
PR 14-MAR-2001; 2001US-0808689.  
PR 22-MAR-2001; 2001US-0816744.  
PR 05-APR-2001; 2001US-0828366.  
PR 10-MAY-2001; 2001US-0854208.  
PR 18-MAY-2001; 2001US-0860216.  
PR 25-MAY-2001; 2001US-0866028.  
PR 25-MAY-2001; 2001US-0866034.  
PR 01-JUN-2001; 2001US-0872035.  
PR 05-JUN-2001; 2001US-0874503.  
PR 14-JUN-2001; 2001US-0882636.  
PR 19-JUN-2001; 2001US-0886342.  
PR 21-JUN-2001; 2001US-0887879.  
PR 18-JUL-2001; 2001US-0908827.  
PR 06-AUG-2001; 2001US-0924419.  
PR 09-AUG-2001; 2001US-0927796.  
PR 16-AUG-2001; 2001US-0931836.  
PR 19-DEC-2001; 2001US-0028072.  
XX  
XX (GETH ) GENENTECH INC.  
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
WPI: 2003-332040/31.  
P-PSDB; ABU66797.  
XX New secreted and transmembrane PRO nucleic acids, useful for gene  
PT therapy, in chromosome and gene mapping, as chromosome markers, in  
PT tissue typing, and in chromosome identification  
XX

PS Claim 2; Fig 455; 560pp; English.

xx The present invention relates to the isolation of novel human PRO polypeptides, and the polynucleotide sequences encoding them. The PRO polypeptides are secreted and transmembrane proteins. The PRO polypeptides are useful for detecting other PRO polypeptides, for linking bioactive molecules to cells expressing PRO polypeptides, for modulating biological activities of cells expressing PRO polypeptides, and for identifying agonists or antagonists.

CC The PRO polypeptides are useful for stimulating the release of tumour necrosis factor (TNF)-alpha from human blood, for stimulating the proliferation or differentiation of chondrocytes, and detecting the presence of tumours. The polynucleotide sequences encoding PRO polypeptides are useful as hybridisation probes, in chromosome and gene mapping, in the generation of antisense RNA and DNA, in the preparation of PRO polypeptides, for generating transgenic animals or knockout animals, for the genetic analysis of individuals with genetic disorders, and in gene therapy. ACA03603-ACA03877 represent cDNAs encoding the human PRO polypeptides of the invention.

CC Note: The sequence data for this patent was obtained in electronic format directly from the USPTO web site at [seqdata.uspto.gov/psipsdIDentry.html](http://seqdata.uspto.gov/psipsdIDentry.html).

xx Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

#### Alignment Scores:

Pred. No.:	2.07e-137	Length:	1570
Score:	1608.00	Matches:	293
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	25	Gaps:	0

US-09-936-271B-14 (1-293) x ACA03830 (1-1570)

Qy	1	MetAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuThrAlaLeuLeu 20
Db	354	ATGGCTACACAGACCCCTGGATGGTGGTCTCTGTCTGATCAGCGCTTGCCT 413
Qy	21	LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db	414	CTGGGGGTACAGACATGTTCTCGCAACAATGATGTTTCTGTGACCACTCTAAC 473
Qy	41	ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAlaArgSer 60
Db	474	ACCGTGCCCTCTGGGAGCAACAGACCTGGGAGCTGGGCGGGGAGACGCCGCTCG 533
Qy	61	AspAspSerSerArgIleIleAsnGlySerAspCysAspMethHisThrGlnProTrp 80
Db	534	GATGACAGCAGCAGCCGATCATATGATGATCCGACTGCGATATGCACCCAGCGTGG 593
Qy	81	GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db	594	CAGCGCGCGTGTGTAAAGCCCAACCACTCTACTCGGGCGGTGTGTGTCATCCA 653
Qy	101	GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120
Db	654	CAGTGGCTGTCTACGGCGCCCACTGCAGAGAAAGATTTTCAGAGTCCGCTCTCGGCCAC 713
Qy	121	TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db	714	TACTCCTGTACACGATTTATGATCTGGGACGACATGTTCCAGGGGGTCAATCCATC 773
Qy	141	ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLeuLysLeuAsn 160
Db	774	CCCCACCTGGCTACTCCACCTGGCCACTCTAACGACCTCATGCTCATCAACTGAAC 833
Qy	161	ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db	834	AGAAGAATTCTGCTCCCACTAAAGATGTGACACCCATCAACGCTCTCTCTATTGCTCCTCT 893
Qy	181	AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200

Db	894	GCTGGGACAAAGTGTCTTGGTGTCTGGCTGGGGGACAAACCAAGAGCCCCCAAGTGCATTC 953
Qy	201	ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db	954	CCTAAGGTCTCTCAGTGTCTGATATATCAGCGTCTAAGTCTAGTAAGGTCGAGGATGCT 1013
Qy	221	TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db	1014	TACCCGAGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTCC 1073
Qy	241	CysGlnClyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db	1074	TGCCAGGGTGATTTCTGGGGGGCTGTGTCTGCAATGGCTCCCTGCGAGGACTCGTGTCC 1133
Qy	261	TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db	1134	TGGGAGATTACCTTGTGCGCGCCCAACAGACCGGGTCTTACAGCAACCTCTGCAAG 1193
Qy	281	PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db	1194	TTCCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1232

RESULT 13  
ACA04251 standard; cDNA; 1570 BP.

XX	ACA04251;
XX	27-MAY-2003 (first entry)
XX	Human cDNA encoding a secreted/transmembrane protein, SEQ ID 455.
XX	Human; ss; gene; secreted protein; transmembrane protein; PRO;
KW	Inflammatory disease; organ failure; atherosclerosis; cardiac injury;
KW	infertility; birth defects; premature aging; AIDS; biosensor;
KW	acquired immunodeficiency syndrome; cancer; diabetic complication;
KW	bioreactor; tumour.
OS	Homo sapiens.
XX	US2003032155-A1.
XX	13-FEB-2003.
XX	03-MAY-2002; 2002US-0137865.
XX	31-MAR-1997; 97WO-US05230.
PR	12-JUN-1998; 98WO-US12456.
PR	14-JUL-1998; 98WO-US14552.
PR	28-AUG-1998; 98WO-US17888.
PR	10-SEP-1998; 98WO-US18824.
PR	14-SEP-1998; 98WO-US19093.
PR	14-SEP-1998; 98WO-US19094.
PR	14-SEP-1998; 98WO-US19177.
PR	16-SEP-1998; 98WO-US19330.
PR	17-SEP-1998; 98WO-US19437.
PR	07-OCT-1998; 98WO-US21141.
PR	29-OCT-1998; 98WO-US22992.
PR	20-NOV-1998; 98WO-US24855.
PR	01-DEC-1998; 98WO-US25108.
PR	08-JAN-1999; 99WO-US00106.
PR	08-MAR-1999; 99WO-US05028.
PR	10-MAR-1999; 99WO-US05190.
PR	20-APR-1999; 99WO-US08615.
PR	14-MAY-1999; 99WO-US10733.
PR	02-JUN-1999; 99WO-US12252.
PR	01-SEP-1999; 99WO-US20111.
PR	08-SEP-1999; 99WO-US20594.
PR	13-SEP-1999; 99WO-US20944.
PR	15-SEP-1999; 99WO-US21090.
PR	13-SEP-1999; 99WO-US21547.
PR	03-OCT-1999; 99WO-US23089.

```
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 22-DEC-1999; 99WO-US30720.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 24-FEB-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05746.
PR 10-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 11-AUG-2000; 2000WO-US22031.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30952.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
PR 01-MAY-2001; 2001WO-US06666.
PR 25-MAY-2001; 2001WO-US17092.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19692.
PR 22-JUN-2001; 2001WO-US20116.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 20-DEC-2000; 2000US-0747259.
PR 28-FEB-2001; 2001US-0796498.
PR 09-MAR-2001; 2001US-0802706.
PR 14-MAR-2001; 2001US-0808689.
PR 22-MAR-2001; 2001US-0816744.
PR 05-APR-2001; 2001US-0828366.
PR 10-MAY-2001; 2001US-0854208.
PR 10-MAY-2001; 2001US-0854280.
PR 18-MAY-2001; 2001US-0860216.
PR 25-MAY-2001; 2001US-0866028.
PR 25-MAY-2001; 2001US-0866034.
PR 01-JUN-2001; 2001US-0872035.
PR 05-JUN-2001; 2001US-0874503.
PR 14-JUN-2001; 2001US-0882636.
PR 19-JUN-2001; 2001US-0886342.
PR 21-JUN-2001; 2001US-0887879.
PR 18-JUL-2001; 2001US-0908827.
PR 06-AUG-2001; 2001US-0924419.
PR 09-AUG-2001; 2001US-0927796.
PR 16-AUG-2001; 2001US-0931836.
PR 19-DEC-2001; 2001US-0028072.
XX

PA (GETH ) GENENTECH INC.
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX WPI; 2003-331925/31.
DR DR P-PSDB; ABU67073.
XX
XX New secreted and transmembrane nucleic acids and polypeptides.
PT designated as PRO, useful for treating inflammation, organ failure,
PT atherosclerosis, cardiac injury, infertility, birth defects, premature
PT aging, AIDS, or cancer
XX
XX Claim 2; Fig 455; 659pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising, or which is
CC at least 80% identical to, or the full-length coding sequence of, any of
CC the 275 nucleotide sequences, encoding the corresponding PRO polypeptide
CC (one of 275 secreted or transmembrane proteins). The nucleic acid
CC further comprises the full-length coding sequence of the DNA deposited
CC under American Type Culture Collection (ATCC) accession number in a list
CC given in the specification. Also included are vectors and host
CC cells for producing PRO proteins, PRO fusion proteins, anti-PRO
CC antibodies, PRO extracellular domains and mature sequences, methods
CC of detecting PRO proteins, methods for stimulating the release of
CC TNF-alpha (tumour necrosis factor alpha) from human blood.
CC (and the proliferation of differentiation of chondrocyte cells, the
CC proliferation of, or gene expression in pericyte cells, the release or
CC proteoglycans from cartilage, proliferation of inner ear utricular
CC supporting cells, the proliferation of T-lymphocyte cells, the release
CC of a cytokine from peripheral blood mononuclear cells (PBMC), or the
CC proliferation of endothelial cells), a method for modulating the uptake
CC of glucose or free fatty acid (FFA) by skeletal muscle cells.
CC a method for inhibiting the binding of A-peptide to factor VIIa,
CC or the differentiation of adipocyte cells, a method for detecting the
CC presence of a tumour in a mammal and an oligonucleotide probe derived
CC from any of the nucleotide sequences cited above. The nucleic acids and
CC polypeptides are useful for treating inflammatory diseases, organ
CC failure, atherosclerosis, cardiac injury, infertility, birth defects,
CC premature aging, AIDS (acquired immunodeficiency syndrome), cancer, or
CC diabetic complications. The nucleic acids are useful as hybridisation
CC probes, in chromosome and gene mapping, and in generating antisense RNA
CC or DNA. The polypeptides are useful as pharmaceuticals, diagnostics,
CC biosensors or bioreactors. Both are useful in tissue typing.
CC The present sequence encodes a PRO protein of the invention.
XX
SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

Alignment Scores:
Pred. No.: 2,07e-137 Length: 1570
Score: 1608.00 Matches: 293
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 25 Gaps: 0

US-09-936-271B-14 (1-293) x ACA04251 (1-1570)
QY 1 MetaLThrAlaArgProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeu 20
|||||
Db 354 ATGGCTACAGCAAGACCCCTGGATGTGGTGTCTGTCTGTATCATCAGCTTGCTT 413
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
|||||
Db 414 CTGGGGGTACAGAGCATGTTCTGCCAACAAATGATGTTCTGTGTACACCCCTTAAC 473
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
|||||
Db 474 ACCGTGCCCTCTGGGGAGCAACACGAGACCTGGGAGCTGGGGCCGGGGAACACGCCCGTCG 533
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMethIleThrGlnProTrp 80
|||||
```

```
Db 534 GATGACAGCAGCGCATCATCAATGGATCGACTGCGATATGCACACCCAGCGTGG 593
Qy 81 GlnAlaLeuLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100
Db 594 CAGCGCGCGTGTGTAGGCCCAACCAAGCTCTACTCGGGCGGTGTGTGGTCATCCA 653
Qy 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120
Db 654 CAGTGGGTGCTACGGCGGCCCACTGCAGGAAGAAAGTTTTCAGAGTCCGTCFCGCCAC 713
Qy 121 TyrSerLeuSerProValTyrGlnSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140
Db 714 TACTCCCTGTCCAGGTATTAATCTGGCGACAGATGTCCAGGGGGTCAATCAATC 773
Qy 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLysLeuAsn 160
Db 774 CCCACCCCTGGCTACTCCCACTCCGCACTCTAACGACCTCATGTCTCACTCAACTGAAC 833
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
Db 834 AGAAGAATTGCTCCCACTAAAGATGTGAGACCATCAACGCTCTCTCATTTGCTCTCT 893
Qy 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200
Db 894 GCTGGGACAAAGTCTTGGTGTCTGGCTGGGGGACAAACCAAGAGCCCCAAGTGCATTC 953
Qy 201 ProlLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
Db 954 CCTAAGGTCTCCAGTCTTGAATATCAGCGTGTCTAAGTCAAGAAAGGTCCGAGGATGCT 1013
Qy 221 TyrProArgGlnIleAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240
Db 1014 TACCCGAGACAGATAGATGACACCATGTTCTGGCGCGGTGACAAAGCAGGTAGAGCTCC 1073
Qy 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260
Db 1074 TGGCAGGGTGATTTCTGGGGGGCTGTGGTCTGCAATGGCTTCCCTGCGAGGACTCGTGTC 1133
Qy 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280
Db 1134 TGGGAGATTAACCTTGTCCCGGCCCAACAGACCGGGGTGCTACAGGAACCTCTGCAAG 1193
Qy 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db 1194 TTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1232

RESULT 14
ABX89368
ID ABX89368 standard; cDNA; 1570 BP.
XX
AC ABX89368;
XX
DT 13-MAY-2003 (first entry)
XX
DE DNA encoding novel secreted and transmembrane protein' PRO1132.
KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
KW cardiac insufficiency disorder; cancer; tumour; immune response;
KW adrenal cortical capillary endothelial growth; c-fos induction;
KW vascular endothelial growth factor inhibition; VEGF inhibition;
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
KW retinal neurons cell survival; rod photoreceptor cell survival;
KW retinal disorder; retinitis pigmentosa; kidney disease;
KW mammalian kidney mesangial cell proliferation; Berger disease;
KW dermatitis; herpetic keratitis; Crohn's disease; chondrocyte proliferation;
KW chondrocyte redifferentiation; sports injury; arthritis; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003017563-A1.
XX
PD 23-JAN-2003.
XX
```

```
PF 07-MAY-2002; 2002US-0140808.
XX
PR 31-MAR-1997; 97WO-US05230.
PR 12-JUN-1998; 98WO-US12456.
PR 14-JUL-1998; 98WO-US14552.
PR 28-AUG-1998; 98WO-US17888.
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19093.
PR 14-SEP-1998; 98WO-US19094.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 07-OCT-1998; 98WO-US21141.
PR 29-OCT-1998; 98WO-US22991.
PR 29-OCT-1998; 98WO-US22992.
PR 20-NOV-1998; 98WO-US24855.
PR 01-DEC-1998; 98WO-US25108.
PR 05-JAN-1999; 99WO-US00106.
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99WO-US05190.
PR 20-APR-1999; 99WO-US08615.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20111.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 22-DEC-1999; 99WO-US30720.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.
PR 11-FEB-2000; 2000WO-US03565.
PR 18-FEB-2000; 2000WO-US04341.
PR 18-FEB-2000; 2000WO-US04342.
PR 22-FEB-2000; 2000WO-US04414.
PR 24-FEB-2000; 2000WO-US04914.
PR 01-MAR-2000; 2000WO-US05004.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05746.
PR 02-MAR-2000; 2000WO-US05841.
PR 10-MAR-2000; 2000WO-US06319.
PR 15-MAR-2000; 2000WO-US06884.
PR 20-MAR-2000; 2000WO-US07377.
PR 21-MAR-2000; 2000WO-US07532.
PR 30-MAR-2000; 2000WO-US08439.
PR 17-MAY-2000; 2000WO-US13705.
PR 22-MAY-2000; 2000WO-US14042.
PR 30-MAY-2000; 2000WO-US14941.
PR 02-JUN-2000; 2000WO-US15264.
PR 28-JUL-2000; 2000WO-US20710.
PR 11-AUG-2000; 2000WO-US22031.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 08-NOV-2000; 2000WO-US30952.
PR 10-NOV-2000; 2000WO-US30873.
PR 01-DEC-2000; 2000WO-US32678.
PR 20-DEC-2000; 2000WO-US34956.
PR 28-FEB-2001; 2001WO-US06520.
```

PR 01-MAR-2001; 2001WO-US06666.  
PR 25-MAY-2001; 2001WO-US17092.  
PR 01-JUN-2001; 2001WO-US17800.  
PR 20-JUN-2001; 2001WO-US19692.  
PR 22-JUN-2001; 2001WO-US20116.  
PR 29-JUN-2001; 2001WO-US21066.  
PR 09-JUL-2001; 2001WO-US21735.  
PR 20-DEC-2000; 2000US-0747259.  
PR 28-FEB-2001; 2001US-0796498.  
PR 09-MAR-2001; 2001US-0802706.  
PR 14-MAR-2001; 2001US-0808689.  
PR 22-MAR-2001; 2001US-0816744.  
PR 05-APR-2001; 2001US-0828366.  
PR 10-MAY-2001; 2001US-0854208.  
PR 18-MAY-2001; 2001US-0854280.  
PR 25-MAY-2001; 2001US-0860216.  
PR 25-MAY-2001; 2001US-0866028.  
PR 25-MAY-2001; 2001US-0866034.  
PR 01-JUN-2001; 2001US-0872035.  
PR 05-JUN-2001; 2001US-0874503.  
PR 14-JUN-2001; 2001US-0882636.  
PR 19-JUN-2001; 2001US-0886342.  
PR 21-JUN-2001; 2001US-0887879.  
PR 18-JUL-2001; 2001US-0908827.  
PR 06-AUG-2001; 2001US-0924419.  
PR 09-AUG-2001; 2001US-0927796.  
PR 16-AUG-2001; 2001US-0931836.  
PR 19-DEC-2001; 2001US-0028072.  
XX

(GETH ) GENENTECH INC.

PA Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;  
PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;  
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;  
XX  
DR WPI; 2003-148238/14.  
DR P-PSDB; ABU59878.  
XX

PT Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346  
PT and PRO1375, which stimulate proliferation of stimulated T-lymphocytes  
PT are therapeutically useful for enhancing immune response and in cancer  
PT treatments

PS Claim 2; Fig 455; 659pp; English.

CC The invention describes an isolated human PRO polypeptide. The PRO  
CC polypeptides are useful in detecting PRO polypeptides in a sample, in  
CC linking a bioactive molecule to a cell expressing a PRO polypeptide, and  
CC in modulating at least one biological activity of a cell expressing a PRO  
CC polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus  
CC useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186  
CC stimulate adrenal cortical capillary endothelial growth, and PRO536,  
CC PRO943, PRO828, PRO846, PRO1068 or PRO535, PRO826, PRO819, PRO1126,  
CC PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus  
CC useful for treating conditions or disorders where angiogenesis would be  
CC beneficial, e.g. wound healing and antagonist of this polypeptide are  
CC useful for treating cancerous tumours. PRO812 inhibits vascular  
CC endothelial growth factor (VEGF) stimulated proliferation of endothelial  
CC cells and is thus useful for inhibiting endothelial cell growth in  
CC mammals which would be beneficial in inhibiting tumour growth. PRO826,  
CC PRO1068, PRO1184, and PRO1346 and PRO1375 stimulate proliferation of  
CC stimulated T-lymphocytes and are therapeutically useful for enhancing  
CC immune response. PRO828, PRO826, PRO1068 or PRO1132 enhance survival of  
CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of  
CC rod photoreceptor cells) and therefore are useful for treating retinal  
CC disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813  
CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,  
CC and therefore are useful for treating kidney disorders associated with  
CC decreased mesangial cell function such as Berger disease or other  
CC nephropathies associated with dermatitis, herpeticiformis or Crohn's  
CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the  
CC proliferation and/or redifferentiation of chondrocytes in culture and  
CC are thus useful for treating sports injuries, and arthritis. This

CC sequence encodes a novel human PRO protein.  
XX  
SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

#### Alignment Scores:

Pred. No.: 2,07e-137 Length: 1570  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 25

US-09-936-271B-14 (1-293) x ABX89368 (1-1570)

QY	1	MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu	20
DB	354	ATGGCTACAGCAAGACCCCTCGATGTGGTGCTCTGTGCTCTGATCACAGCTTGCTT	413
QY	21	LeuGlyValThrGluHisValLeuAlaAsnAsnAspValSerCysAspHisProSerAsn	40
DB	414	CTGGGGTTCACAGAGCATGTTCTGCCCAACAATGATGTTCTGTGACCACTCTCTAAC	473
QY	41	ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer	60
DB	474	ACCGTCCCTCTGGGAGCAACCAAGGACTGGGAGCTGGGCGCGGGAAGACGCCGCTCG	533
QY	61	AspAspSerSerArgIleIleAsnGlySerAspCysAspMethHisThrGlnProTrp	80
DB	534	GATGACAGCAGCAGCCGATCATCAATGATGATGATGATGATGATGATGATGATGATG	593
QY	81	GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrcysGlyAlaValLeuValHisPro	100
DB	594	CAGGCGCGCTGTGCTAAGGCCCAACCACTACTGCGGGCGGTGTGGTGATCCCA	653
QY	101	GlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHis	120
DB	654	CAGTGGCTGCTCACGGCGCCCACTGCAGGAGAGAAAGTTTTCAGAGTCCGCTCGGCC	713
QY	121	TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle	140
DB	714	TACTCCCTGTACAGTTTATGATCTGGGACAGATGTTCCAGGGGGTCAATCATC	773
QY	141	ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn	160
DB	774	CCCCACCTGGCTACTCCCACTGGCCACTCTAACGACCTCATGCTCATCAACTGAAC	833
QY	161	ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer	180
DB	834	AGAAGAATTCTGCCCACTAAAGATGTCAGACCCATCAACGCTCTCTCTCATTTGCC	893
QY	181	AlaGlyThrLysCysLeuValSerGlyTyrGlyThrLysSerProGlnValHisPhe	200
DB	894	GCTGGGACAAAGTGTGTGGTGTCTGGCTGGGGACCAACCAAGAGCCCCCAAGTGC	953
QY	201	ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla	220
DB	954	CCTAAGTCTCTCCAGTCTTGAATATCAGCGTCTAAGTCAGAAAAGTCTCGAGGAT	1013
QY	221	TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer	240
DB	1014	TACCCGACAGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAC	1073
QY	241	CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer	260
DB	1074	TGCCAGGGGTATTTCTGGGGGCTGTGGTCTTGAATGGCTCTCCCTGAGGAGCTCG	1133
QY	261	TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys	280
DB	1134	TGGGGAGATTACCTTGTGCCCCGCCCAACAGACCCGGGTGTCTACAGAACCTCTG	1193
QY	281	PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer	293
DB	1194	TTCAACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC	1232

RESULT 15  
ID ABX80318  
XX ABX80318 standard; DNA; 1570 BP.  
AC ABX80318;  
XX  
DT 28-APR-2003 (first entry)  
XX  
DE Human secreted or transmembrane protein related PCR primer #56.  
XX  
KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;  
KW cardiac insufficiency disorder; cancer; tumour; immune response;  
KW adrenal cortical capillary endothelial growth; c-fos induction;  
KW vascular endothelial growth factor inhibition; VEGF inhibition;  
KW endothelial cell growth inhibitor; T-lymphocytes stimulation;  
KW retinal neurons cell survival; rod photoreceptor cell survival;  
KW retinal disorder; retinitis pigmentosa; kidney disorder;  
KW mammalian kidney mesangial cell proliferation; Berger disease;  
KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;  
KW chondrocyte redifferentiation; sports injury; arthritis; PCR;  
KW primer; ss.  
XX  
OS Homo sapiens.  
XX  
PN US2002132252-A1.  
XX  
PD 19-SEP-2002.  
XX  
PF 14-NOV-2001; 2001US-0990442.  
XX  
XX 05-NOV-1997; 97WO-US200069.  
PR 16-SEP-1998; 98WO-US19330.  
PR 17-SEP-1998; 98WO-US19437.  
PR 07-OCT-1998; 98WO-US21141.  
PR 01-DEC-1998; 98WO-US25108.  
PR 05-JAN-1999; 99WO-US00106.  
PR 08-MAR-1999; 99WO-US05028.  
PR 02-JUN-1999; 99WO-US12252.  
PR 15-SEP-1999; 99WO-US21090.  
PR 15-SEP-1999; 99WO-US21547.  
PR 30-NOV-1999; 99WO-US28313.  
PR 01-DEC-1999; 99WO-US28301.  
PR 01-DEC-1999; 99WO-US28634.  
PR 16-DEC-1999; 99WO-US30095.  
PR 20-DEC-1999; 99WO-US30911.  
PR 06-JAN-2000; 2000WO-US00219.  
PR 06-JAN-2000; 2000WO-US00376.  
PR 11-FEB-2000; 2000WO-US03565.  
PR 18-FEB-2000; 2000WO-US04341.  
PR 22-FEB-2000; 2000WO-US04414.  
PR 24-FEB-2000; 2000WO-US04914.  
PR 24-FEB-2000; 2000WO-US05004.  
PR 02-MAR-2000; 2000WO-US05841.  
PR 10-MAR-2000; 2000WO-US06319.  
PR 15-MAR-2000; 2000WO-US06884.  
PR 20-MAR-2000; 2000WO-US07377.  
PR 30-MAR-2000; 2000WO-US08439.  
PR 15-MAY-2000; 2000WO-US13358.  
PR 17-MAY-2000; 2000WO-US13705.  
PR 22-MAY-2000; 2000WO-US14042.  
PR 30-MAY-2000; 2000WO-US14941.  
PR 02-JUN-2000; 2000WO-US15264.  
PR 28-JUL-2000; 2000WO-US20710.  
PR 11-AUG-2000; 2000WO-US22031.  
PR 23-AUG-2000; 2000WO-US23522.  
PR 24-AUG-2000; 2000WO-US23328.  
PR 08-NOV-2000; 2000WO-US30952.  
PR 01-DEC-2000; 2000WO-US32678.  
PR 28-FEB-2001; 2001WO-US06520.  
PR 01-JUN-2001; 2001WO-US17800.  
PR 20-JUN-2001; 2001WO-US19692.  
PR 29-JUN-2001; 2001WO-US21066.  
PR 09-JUL-2001; 2001WO-US21735.  
PR 16-JUN-1997; 97US-049787P.  
PR 17-OCT-1997; 97US-062250P.  
PR 12-NOV-1997; 97US-065186P.  
PR 13-NOV-1997; 97US-065311P.  
PR 24-NOV-1997; 97US-066770P.  
PR 25-FEB-1998; 98US-075945P.  
PR 20-MAR-1998; 98US-078910P.  
PR 28-APR-1998; 98US-083322P.  
PR 07-MAY-1998; 98US-084600P.  
PR 28-MAY-1998; 98US-087106P.  
PR 02-JUN-1998; 98US-087607P.  
PR 02-JUN-1998; 98US-087609P.  
PR 02-JUN-1998; 98US-087759P.  
PR 03-JUN-1998; 98US-087827P.  
PR 04-JUN-1998; 98US-088021P.  
PR 04-JUN-1998; 98US-088023P.  
PR 04-JUN-1998; 98US-088026P.  
PR 04-JUN-1998; 98US-088028P.  
PR 04-JUN-1998; 98US-088029P.  
PR 04-JUN-1998; 98US-088030P.  
PR 04-JUN-1998; 98US-088033P.  
PR 04-JUN-1998; 98US-088326P.  
PR 05-JUN-1998; 98US-088167P.  
PR 05-JUN-1998; 98US-088202P.  
PR 05-JUN-1998; 98US-088212P.  
PR 05-JUN-1998; 98US-088217P.  
PR 09-JUN-1998; 98US-088655P.  
PR 10-JUN-1998; 98US-088734P.  
PR 10-JUN-1998; 98US-088738P.  
PR 10-JUN-1998; 98US-088742P.  
PR 10-JUN-1998; 98US-088810P.  
PR 10-JUN-1998; 98US-088824P.  
PR 10-JUN-1998; 98US-088826P.  
PR 11-JUN-1998; 98US-088858P.  
PR 11-JUN-1998; 98US-088861P.  
PR 11-JUN-1998; 98US-088876P.  
PR 12-JUN-1998; 98US-089105P.  
PR 16-JUN-1998; 98US-089440P.  
PR 16-JUN-1998; 98US-089512P.  
PR 16-JUN-1998; 98US-089514P.  
PR 17-JUN-1998; 98US-089532P.  
PR 17-JUN-1998; 98US-089538P.  
PR 17-JUN-1998; 98US-089598P.  
PR 17-JUN-1998; 98US-089599P.  
PR 17-JUN-1998; 98US-089600P.  
PR 17-JUN-1998; 98US-089653P.  
PR 18-JUN-1998; 98US-089801P.  
PR 18-JUN-1998; 98US-089907P.  
PR 18-JUN-1998; 98US-089908P.  
PR 28-AUG-2001; 2001US-0941992.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
XX Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;  
PI Ferrara N, Fong S, Gerber H, Gertitsen ME, Goddard A, Godowski PJ;  
PI Grimaldi JC, Gurney AL, Kljavin IJ, Napier MA, Pan J, Paoni NF;  
PI Roy MA, Stewart TA, Tumas D, Watanabe CK, Williams PM, Wood WI;  
PI Zhang Z;  
XX  
XX WPI; 2003-247083/24.  
XX  
XX Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346  
PT and PRO1375, which stimulate proliferation of stimulated T-lymphocytes  
PT are therapeutically useful for enhancing immune response and in cancer  
PT treatments  
XX  
XX Example 99; Page 259; 648pp; English.  
XX  
XX The invention describes an isolated human PRO polypeptide. The PRO  
CC polypeptides are useful in detecting PRO polypeptides in a sample, in  
CC linking a bioactive molecule to a cell expressing a PRO polypeptide, and  
CC in modulating at least one biological activity of a cell expressing a PRO

CC polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus  
CC useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186  
CC stimulate adrenal cortical capillary endothelial growth, and PRO536,  
CC PRO943, PRO828, PRO826, PRO1068 or PRO535, PRO826, PRO819, PRO1126,  
CC PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus  
CC useful for treating conditions or disorders where angiogenesis would be  
CC beneficial, e.g. wound healing and antagonism of this polypeptide are  
CC useful for treating cancerous tumours. PRO812 inhibits vascular  
CC endothelial growth factor (VEGF) stimulated proliferation of endothelial  
CC cells and is thus useful for inhibiting endothelial cell growth in  
CC mammals which would be beneficial in inhibiting tumour growth. PRO826,  
CC PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of  
CC stimulated T-lymphocytes and are therapeutically useful for enhancing  
CC immune response. PRO828, PRO826, PRO1068 or PRO132 enhance survival of  
CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of  
CC rod photoreceptor cells) and therefore are useful for treating retinal  
CC disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813  
CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,  
CC and therefore are useful for treating kidney disorders associated with  
CC decreased mesangial cell function such as Berger disease or Crohn's  
CC nephropathies associated with dermatitis, herpeticiformis or Crohn's  
CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the  
CC proliferation and/or redifferentiation of chondrocytes in culture and  
CC are thus useful for treating sports injuries, and arthritis. This  
CC sequence represents a primer used in the isolation of DNA encoding  
CC novel human PRO polypeptides.  
XX  
SQ Sequence 1570 BP; 332 A; 485 C; 424 G; 329 T; 0 other;

Alignment Scores:

Pred. No.: 2,07e-137 Length: 1570  
Score: 1608.00 Matches: 293  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 25 Gaps: 0

US-09-936-271b-14 (1-293) x ABX80318 (1-1570)

QY 1 MetAlaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
DB 354 ATGGCTACAGCAGACAGACCCCTGGGTGGGTCTGTGCTGATCAGCCCTTGCTT 413  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
DB 414 CTGGGGTTCACAGAGCATGTTCTGCCCAACATGATGTTCTGTGACACCCCTTAAC 473  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
DB 474 ACCGTGCCCTCTGGGAGCAACAGGACCTGGGAGCTGGGGCGGGGAAGACGCCCGTGG 533  
QY 61 AspAspSerSerArgIleIleAsnGlySerAspCysAspMethIleThrGlnProTrp 80  
DB 534 GATGACGAGCAGCGCGCATCATCAATGATCGGACTGGCATGACACCCAGCGGTGG 593  
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
DB 594 CAGCGCGCGCTGTGCTAGGCCCAACACAGCTCTACTGGGGCGGTGTGGTGATCCA 653  
QY 101 GlnTrpLeuLeuThrAlaAlaHisCysArgLysValPheArgValArgLeuGlyHis 120  
DB 654 CAGTGGCTGCTCAGCGCGCCCTGCAGGAAGAAAGTTTTCAGAGTCCGTCTCGGCCAC 713  
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIle 140  
DB 714 TACTCCCTGTACCAAGTTTATGAATCTGGGCAGCAGATGTTCCAGGGGTCAAATCCATC 773  
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
DB 774 CCCACCCCTGGTACTCCACCCCTGGCCACTCTAACGACCTCATGCTCAAACTGAAC 833  
QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
DB 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180

DB 834 AGAAGAAATTCGTCCTCCACTAAAGATGTGACAGCCCATCAACGTCTCTCTCATTTGCCCTCT 893  
QY 181 AlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPhe 200  
DB 894 GCTGGGACAAAGTGTGTTGGTCTGGCTGGGGGACAAACCAAGAGCCCCCAAGTGCACATTC 953  
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
DB 954 CCTAAGGTCTCCAGTGTGAATATCAGCGTCTAAGTCAGAAAAAGGTGCGAGGATGCT 1013  
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSer 240  
DB 1014 TACCCGAGACAGATAGATGACCATGTTCTGCGCGGTGACAAAAGCAGGTAGAGACTCC 1073  
QY 241 CysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSer 260  
DB 1074 TGGCAGGGTGATCTCGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGAGACTCGTGTC 1133  
QY 261 TrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLys 280  
DB 1134 TGGGGAGATTACCTTTGTCGCCGCCAACAGACCGGGTGTCTACAGAACTCTTGCAG 1193  
QY 281 PheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293  
DB 1194 TTCACCAAGTGGATCCAGGAAACCATCCAGGCCAACTCC 1232

Search completed: August 5, 2003, 12:30:49  
Job time : 282 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: August 5, 2003, 12:11:15 ; Search time 77 Seconds

(without alignments)  
1679.549 Million cell updates/sec

Title: US-09-936-271B-14

Perfect score: 1608

Sequence: 1 MATARPPWVWLICALITALL.....VTNLCKFKTQWQETIQANS 293

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool/US09936271/runat\_04082003\_104246\_25727/app\_query.fasta\_1.455  
-DB=Issued Patents NA -Qfmt=fastap -SUFFIX=rni -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09936271@cgn\_1.1\_48 -runat\_04082003\_104246\_25727 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCUTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1608	100.0	1570	4	US-09-996-243-308
2	1595	99.2	1476	2	Sequence 308, App
3	1595	99.2	1476	3	Sequence 2, Appli
4	1595	99.2	1476	4	Sequence 2, Appli
5	1403	87.3	1504	4	Sequence 1, Appli
6	725	45.1	735	4	Sequence 94, Appl
7	671	41.7	765	4	Sequence 524, App
8	654.5	40.7	1052	4	Sequence 10, Appl
9	652.5	40.6	1146	4	Sequence 247, App
10	652.5	40.6	1166	3	Sequence 7, Appli
11	650.5	40.5	833	2	Sequence 2, Appli
12	650.5	40.5	1192	3	Sequence 8, Appli

13	650.5	40.5	1292	4	US-09-205-258-189	Sequence 189, Appl
14	650.5	40.5	1314	3	US-09-025-059-2	Sequence 2, Appl
15	639.5	39.8	994	3	US-09-008-271A-19	Sequence 19, Appl
16	638.5	39.7	944	3	US-09-070-526-1	Sequence 1, Appl
17	638	39.7	871	1	US-08-744-026-2	Sequence 2, Appl
18	638	39.7	871	2	US-09-102-732-2	Sequence 2, Appl
19	638	39.7	871	3	US-09-261-767-2	Sequence 7, Appl
20	638	39.7	871	4	US-08-969-987-7	Sequence 7, Appl
21	638	39.7	1248	3	US-09-020-956-171	Sequence 171, App
22	638	39.7	1248	3	US-09-030-607-171	Sequence 171, App
23	638	39.7	1248	4	US-09-439-313-171	Sequence 171, App
24	638	39.7	1248	4	US-09-352-616A-171	Sequence 171, App
25	638	39.7	1248	4	US-09-232-149A-171	Sequence 171, App
26	635	39.5	1037	4	US-09-386-642-60	Sequence 60, Appl
27	632.5	39.3	1049	4	US-09-386-642-9	Sequence 9, Appl
28	620	38.6	1214	4	US-09-439-313-225	Sequence 225, App
29	620	38.6	1214	4	US-09-352-616A-225	Sequence 225, App
30	620	38.6	1214	4	US-09-232-149A-225	Sequence 225, App
31	620	38.6	1215	4	US-09-439-313-326	Sequence 326, App
32	620	38.6	1215	4	US-09-352-616A-326	Sequence 326, App
33	620	38.6	1215	4	US-09-232-149A-326	Sequence 326, App
C 34	612.5	38.1	969	3	US-09-502-600-30	Sequence 30, Appl
35	612.5	38.1	986	2	US-08-557-146-1	Sequence 1, Appl
36	612.5	38.1	986	2	US-09-154-344-1	Sequence 1, Appl
C 37	612.5	38.1	1089	3	US-08-930-188-1	Sequence 1, Appl
C 38	612.5	38.1	1089	3	US-08-930-188-3	Sequence 3, Appl
C 39	612.5	38.1	1089	5	PCT-US96-04294-1	Sequence 1, Appl
C 40	612.5	38.1	1089	5	PCT-US96-04294-3	Sequence 3, Appl
41	590	36.7	1386	2	US-08-897-340-1	Sequence 1, Appl
42	590	36.7	1386	3	US-09-252-329-1	Sequence 1, Appl
C 43	585	36.4	1364	4	US-09-280-116-20	Sequence 20, Appl
44	582.5	36.2	1265	3	US-09-020-956-173	Sequence 173, App
45	582.5	36.2	1265	3	US-09-030-607-173	Sequence 173, App

ALIGNMENTS

RESULT 1  
US-09-996-243-308  
; Sequence 308, Application US/09996243  
; Patent No. 6478825  
; GENERAL INFORMATION:  
; APPLICANT: Ashkenazi, Avi J.  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan L.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gerber, Hanspeter  
; APPLICANT: Gerritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, J. Christopher  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Kljavin, Ivar J.  
; APPLICANT: Napier, Mary A.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; APPLICANT: Roy, Margaret Ann  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K.  
; APPLICANT: Williams, P. Mickey  
; APPLICANT: Wood, William I.  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic Acids Encoding the Same  
; FILE REFERENCE: P2730P1C13  
; CURRENT APPLICATION NUMBER: US/09/996,243  
; CURRENT FILING DATE: 2001-11-14  
; PRIOR APPLICATION NUMBER: 60/049787  
; PRIOR FILING DATE: 1997-06-16



US-09-936-271B-14 (1-293) x US-08-824-874-2 (1-1476)

```
Oy 1 MetaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuLeuThrAlaLeuLeu 20
Db 277 ATGGCTACAGACAGACCCCTGGATGTGGTGTCTGTCTGTATCAGACCTTGCTT 336
Oy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 337 CTGGGGTCTACAGAGCATGTCTCGCAACAATGATTTCTGTGACCACTTCAAC 396
Oy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db 397 ACCGTGCCCTCTGGGACACACAGGACCTGGAGCTGGGGCCGGGAGAGCCCGGTGC 456
Oy 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 457 GATGACAGCAGCAGCCGCATCATCAATGGATCCGACTCGCATATGCACACCCAGCCGTG 516
Oy 81 -GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrcysGlyAlaValLeuValHisPr 100
Db 517 GCAGGCGCGCTGTGTCTAAGGCCCCAACCACTCTACTCGGGGGCGGTGTGGTGCATCC 576
Oy 100 oGlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHI 120
Db 577 ACAGTGGCTGCTCAGCGCCGCCACTGCAGGAAGAATTTTCAGACTCCGCTCGGCCA 636
Oy 120 sTyrSerLeuSerProValTyrGluSerGlyGlnMetPheGlnGlyValLysSerIl 140
Db 637 CTACTCCCTGTCCACAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGTCAAAATCCAT 696
Oy 140 eProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLysLeuAs 160
Db 697 CCCCCACCTGGCTACTCCACCTGGCCACTCTAAGCAGCTCTAAGTCTCATCAACACTGAA 756
Oy 160 nArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProse 180
Db 757 CAGAAAGATTGCTCCCACTAAGATGTCAGACCATCAACGTCTCTCTATTGCTCCTC 816
Oy 180 rAlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPh 200
Db 817 TGCTGGGACAAAGTGTGTGTGCTGGGCTGGGGGACCAACCAAGAGCCCAAGTGCATT 876
Oy 200 eProLysValLeuGlnCysLeuAsnIleSerValLeuSerClnLysArgCysGluAspAl 220
Db 877 CCCTAAGGTCTCCAGTGTCTGAATATACAGCGTGTCTAAGTACAGAAAGGTGCGAGATGC 936
Oy 220 aTyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspse 240
Db 937 TTACCCGAGACAGATAGATGACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTC 996
Oy 240 rCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValse 260
Db 997 CTGCCAGGGTGATTCTGGGGGCGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGT 1056
Oy 260 rTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysLy 280
Db 1057 CTGGGGAGATTACCTCTGTGCGCGGCCCAACAGACCGGGTGTCTACAGAACTCTGCAA 1116
Oy 280 sPheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db 1117 GTTCACCAAGTGGATCCAGGAACCATCCAGGCCCAACTCC 1156
```

## RESULT 3

```
US-09-210-084-2
; Sequence 2, Application US/09210084
; Patent No. 6197511
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; TITLE OF INVENTION: NOVEL KALLIKREIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
```

```
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/210,084
; FILING DATE:
; CLASSIFICATION:
```

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/824,874

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0252 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1476 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

LIBRARY: KERANOT02

CLONE: 820694

US-09-210-084-2

## Alignment Scores:

```
Pred. No.: 2,12e-159 Length: 1476
Score: 1595.00 Matches: 293
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 0
Query Match: 99.19% Indels: 1
DB: 3 Gaps: 0
```

US-09-936-271b-14 (1-293) x US-09-210-084-2 (1-1476)

```
Oy 1 MetaThrAlaArgProProTrpMetTrpValLeuCysAlaLeuLeuThrAlaLeuLeu 20
Db 277 ATGGCTACAGACAGACCCCTGGATGTGGTGTCTGTCTGTATCAGACCTTGCTT 336
Oy 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 337 CTGGGGTCTACAGAGCATGTCTCGCAACAATGATTTCTGTGACCACTTCAAC 396
Oy 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db 397 ACCGTGCCCTCTGGGACACACAGGACCTGGAGCTGGGGCCGGGAGAGCCCGGTGC 456
Oy 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 457 GATGACAGCAGCAGCCGCATCATCAATGGATCCGACTCGCATATGCACACCCAGCCGTG 516
Oy 81 -GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrcysGlyAlaValLeuValHisPr 100
Db 517 GCAGGCGCGCTGTGTCTAAGGCCCCAACCACTCTACTCGGGGGCGGTGTGGTGCATCC 576
Oy 100 oGlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHI 120
Db 577 ACAGTGGCTGCTCAGCGCCGCCACTGCAGGAAGAATTTTCAGACTCCGCTCGGCCA 636
Oy 120 sTyrSerLeuSerProValTyrGluSerGlyGlnMetPheGlnGlyValLysSerIl 140
Db 637 CTACTCCCTGTCCACAGTTTATGAATCTGGCAGCAGATGTTCCAGGGGTCAAAATCCAT 696
Oy 140 eProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLysLeuAs 160
```

```
Db 697 CCCCCACCCCTGGCTACTCCACCCCTGGCCACTCTAACGACCTCATCTCAAACTGAA 756
QY 160 nArgArgIleArgProThrLysAspValArgProIleAsnValSerHisCysProse 180
Db 757 CAGAGAATTCGTCCACCTAAAGATGTGACAGCCCATCAACGCTCTCTCTCATTTGCTCCTC 816
QY 180 rAlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPh 200
Db 817 TCGTGGACAAAGTCTGTGGTCTGGCTGGGGGACAAACAGAGCCCAAGTGCACAT 876
QY 200 eProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAl 220
Db 877 CCCTAAGTCTCCAGTCTTGAATATACGCTGCTAAGTCAGAAAAGGTGCGAGGATGC 936
QY 220 aTyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSe 240
Db 937 TTACCCGACAGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTC 996
QY 240 rCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSe 260
Db 997 CTGCCAGGCTGATCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCT 1056
QY 260 rTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysly 280
Db 1057 CTGGGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACCAACTCTGCAA 1116
QY 280 sPheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
Db 1117 GTTCACCAAGTGGATCCAGGAACCATCCAGGCCAACTCC 1156
```

## RESULT 4

US-09-764-762-2

Sequence 2, Application US/09764762

Patent No. 6472195

GENERAL INFORMATION:

APPLICANT: Hillman, Jennifer L.

Lal, Preeti

TITLE OF INVENTION: NOVEL KALLIKREIN

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/764,762

FILING DATE: 16-Jan-2001

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/210,084

FILING DATE: &lt;Unknown&gt;

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J.

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0252 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1476 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

```
; LIBRARY: KERANOT02
; CLONE: 820694
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-764-762-2
```

## Alignment Scores:

```
Pred. No.: 2,12e-159 Length: 1476
Score: 1595.00 Matches: 293
Percent Similarity: 99.66% Conservative: 0
Best Local Similarity: 99.66% Mismatches: 0
Query Match: 99.19% Indels: 1
DB: 4 Gaps: 0
```

US-09-936-271b-14 (1-293) x US-09-764-762-2 (1-1476)

```
QY 1 MetAlaThrAlaArgProProThrMetTrpValLeuCysAlaLeuIleThrAlaLeu 20
Db 277 ATGGCTACAGCAAGACCCCTGGATGTGGGTGCTCTGTCTGTATCAGAGCTTGCTT 336
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40
Db 337 CTGGGGGTACAGAGCATGTTCTCGCCACAATGATGTTTCTGTGACCACTCTTAAC 396
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60
Db 397 ACCGTGCCCTCTGGGAGCAACACGAGCCTGGGAGCTGGGCGCGGGAAGACGCCGCTG 456
QY 61 AspAspSerSerSerArgIleIleAsnGlySerAspCysAspMetHisThrGlnProTrp 80
Db 457 GATGACAGCAGCAGCGCATCATCAATGGATCGGACTGCGATATGATGACACACCCAGC 516
QY 81 -GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPr 100
Db 517 GCAGGCGCGCTGTGTCTAAGGCCCAACAGCTCTACTGCGGGCGCGGTGTGGTGTATCC 576
QY 100 ogInTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHl 120
Db 577 ACAGTGGCTGCTCACGCGCGCCACTGCAGGAAGAAAGTTTTCAGAGTCCGTCTCGCCA 636
QY 120 sTyrSerLeuSerProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIl 140
Db 637 CTACTCCCTGTCCACGATTTATGAATCTGGGACAGAGATGTTCCAGGGGGTCAAACTCAT 696
QY 140 eProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuLysLeuAs 160
Db 697 CCCCCACCCCTGGCTACTCCACCTCGCCACTTAACGACCTCATCTCATCAAACTGAA 756
QY 160 nArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProse 180
Db 757 CAGAGAATTCGTCCACCTAAAGATGTCAAGACCCATCAACGCTCTCTCTCATTTGCCCTC 816
QY 180 rAlaGlyThrLysCysLeuValSerGlyTyrGlyThrThrLysSerProGlnValHisPh 200
Db 817 TCGTGGACAAAGTGTCTGTGCTGCGGCTGGGGGACAAACAGAGCCCAAGTGCACAT 876
QY 200 eProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAl 220
Db 877 CCCTAAGTCTCCAGTCTTGAATATACGCTGCTAAGTCAGAAAAGGTGCGAGGATGC 936
QY 220 aTyrProArgGlnIleAspAspThrMetPheCysAlaGlyAspLysAlaGlyArgAspSe 240
Db 937 TTACCCGACAGACAGATAGATGACACCATGTTCTGCGCGGTGACAAAGCAGGTAGAGACTC 996
QY 240 rCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnGlyLeuValSe 260
Db 997 CTGCCAGGCTGATTTCTGGGGGGCTGTGGTCTGCAATGGCTCCCTGCAGGGACTCGTGTCT 1056
QY 260 rTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCysly 280
Db 1057 CTGGGAGATTACCTTGTGCGCGGCCCAACAGACCGGGTGTCTACACCAACTCTGCAA 1116
QY 280 sPheThrLysTrpIleGlnGluThrIleGlnAlaAsnSer 293
```



QY 284 TrrpGlnGluThrIleGlnAlaAsnSer 293  
|||||  
Db 374 TGGATCCAGGAACCATCCAGGCCAACTCC 345

## RESULT 7

US-09-439-313-524  
; Sequence 524, Application US/09439313  
; Patent No. 6329505  
; GENERAL INFORMATION:  
; APPLICANT: Xu, Jiangchun  
; APPLICANT: Dillon, Davin C.  
; APPLICANT: Mitcham, Jennifer L.  
; APPLICANT: Harlocker, Susan Louise  
; APPLICANT: Jiang Yuqi  
; APPLICANT: Reed, Steven G.  
; APPLICANT: Kalos, Michael  
; APPLICANT: Fanger, Gary  
; APPLICANT: Retter, Mark  
; APPLICANT: Solk, John  
; APPLICANT: Day, Craig  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER  
; FILE REFERENCE: 210121.427C9  
; CURRENT APPLICATION NUMBER: US/09/439,313  
; CURRENT FILING DATE: 1999-11-12  
; NUMBER OF SEQ ID NOS: 575  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 524  
; LENGTH: 765  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-439-313-524

Alignment Scores:  
Pred. No.: 6,11e-62 Length: 765  
Score: 671.00 Matches: 134  
Percent Similarity: 61.77% Conservative: 47  
Best Local Similarity: 45.73% Mismatches: 72  
Query Match: 41.73% Indels: 40  
DB: 4 Gaps: 6

US-09-936-271B-14 (1-293) x US-09-439-313-524 (1-765)

QY 1 MetAlaThrAlaArgProTrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeu 20  
|||||  
Db 1 ATGCCACAGCAAGAAATCCCTGGGCTGGTCTCTGGGGTACCTCATC----- 48  
QY 21 LeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisProSerAsn 40  
|||||  
Db 49 CTGGGTGTC-----GCAGGA 63  
QY 41 ThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSer 60  
: : : : :  
Db 64 TCGCTGCTCTCTGCT----- 78  
QY 61 AspAspSerSerSerArgIleLeuAsnGlySerAspCysAspMethIsthrGlnProTrp 80  
: : : : :  
Db 79 -----AGTGCAGCCAAATCATAAACGGCGAGAGACTGCAGCCCGGCTGCAGCCCTGG 132  
QY 81 GlnAlaAlaLeuLeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisPro 100  
: : : : :  
Db 133 CAGCGGCGACTGTCATGGNA-----AACGAATTTCTCTCGGGCGTCTGGTGATCCG 189  
QY 101 GlnTrpLeuThrAlaAlaHisCysArgLysLysValPheArgValArgLysHis 120  
: : : : :  
Db 190 CAGTGGTGTCTGCAGCCGCACACTGTTCCAGAACTCTACACCATCGGCTGGCGCTG 249  
QY 121 TyrSerLeuSerProValTyrGluSerGlyGlnMetPheGlnGlyValLysSerIle 140  
: : : : :  
Db 250 CACAGTCTTGAGGCCCAACAGAGCCAGGAGCAGATGTTGGAGCCAGCCTCTCCGTA 309  
QY 141 ProHisProGlyTyrSerHisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160

Db 310 CGGCACCCAGAGTACACACACCTTCTCGCTAACGACCTCATGCTCATCAAGTTGGAC 369  
: : : : :  
QY 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180  
: : : : :  
Db 370 GAATCCGTGTCGAGTCTGACACCATCCGAGCATCAGCATTTGCTTCGACGTGCCYACC 429  
: : : : :  
QY 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200  
: : : : :  
Db 430 GCGGGGAACCTCTTGGCTCGCTTCTGGCTGGGGTCTGCTGGCAACGCGCAAGTG----- 483  
QY 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
: : : : :  
Db 484 CTTACCGTGTGTCAGTCCGTGAACGCTCGTGGTGTCTGAGGAGGTCTGCAAGACTC 543  
: : : : :  
QY 221 TyrProArgGlnIleAspAspThrMetPheCysAla---GlyAspLysAlaGlyArgAsp 239  
: : : : :  
Db 544 TATGACCCGCTGTATCACCCACCCAGCATGTTCTGGCGCGCGGAGGCAAGACCAAGAGGAC 603  
: : : : :  
QY 240 SerCysGlnGlyAspSerGlyGlyProValValCysAsnGlySerLeuGlnVal 259  
: : : : :  
Db 604 TCTGCAACGGTGACTCTGGGGGGCCCTGATCTGCAACGGGTACTTGCAGGGCTTGTG 663  
: : : : :  
QY 260 SerTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCys 279  
: : : : :  
Db 664 TCTTTCGAAAAGCCCGCTGTGCCAAGTTGGCGTGCAGGTGTCTACACCACTCTGC 723  
: : : : :  
QY 280 LysPheThrLysTrpIleGlnGluThrIleGlnAlaAsn 292  
: : : : :  
Db 724 AAATTCAGTGGTAGAGAGAAACCGTCCAGGCCAGT 762

## RESULT 8

US-09-386-642-10  
; Sequence 10, Application US/09386642  
; Patent No. 6420157  
; GENERAL INFORMATION:  
; APPLICANT: Darrow, Andrew  
; APPLICANT: Qi, Jensen  
; APPLICANT: Andrade-Gordon, Patricia  
; TITLE OF INVENTION: Zymogen Activation System  
; FILE REFERENCE: ORT-1028  
; CURRENT APPLICATION NUMBER: US/09/386,642  
; CURRENT FILING DATE: 1999-08-31  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 1052  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Fusion gene  
; OTHER INFORMATION: with homo sapien serine protease catalytic domain  
US-09-386-642-10

Alignment Scores:  
Pred. No.: 5,54e-60 Length: 1052  
Score: 654.50 Matches: 129  
Percent Similarity: 62.99% Conservative: 48  
Best Local Similarity: 45.91% Mismatches: 87  
Query Match: 40.70% Indels: 17  
DB: 4 Gaps: 7

US-09-936-271B-14 (1-293) x US-09-386-642-10 (1-1052)

QY 19 LeuLeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAspHisPro 38  
: : : : :  
Db 49 CTCTGCTGCTGGTGTCAATCTACTCTTGTGCCAGGCTGGTCTCGACTACAAG 108  
: : : : :  
QY 39 SerAsnThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyGluAspAla 58  
: : : : :  
Db 109 -----GACGAGCAGCAGCTGGAGCGGCGCT-----CTTGCT 141  
: : : : :  
QY 59 ArgSerAspAspSerSerArgIleIleAsnGlySerAspCys---AspMethIsthr 77





```
Query Match: 40.58% Indels: 7
DB: 3 Gaps: 4

US-09-936-271B-14 (1-293) x US-08-944-483-7 (1-1166)

Qy 65 SerArgIleLeuAsnGlySerAspCysAspMetHisThrGlnProTrpGlnAlaLeu 84
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 160 ACCAGGATCATCAAGAGGGTTCGAGTGCNAGCCTCACTCCAGGCGCTGGCAGGCGCTG 219
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 85 LeuLeuArgProAsnGlnLeuTyrcysGlyAlaValLeuValHisProGlnTrpLeu 104
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 220 TTCAGAGAR---ACCGCGCTACTCTGTGGGGCGAGCTCATCGCCCGCATGGCTCTG 276
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 105 ThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHisTrpSerLeu 124
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 277 ACAGCAGCCCACTCCCTCAAGCCCGCTACATAGTTCACCTGGCGGAGCACAACCTCC 336
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 125 ProValTyrcysGlyGlnGlnMetPheGlnGlyValLysSerIleProHisProGly 144
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 337 AAG---GAGGAGGCTGTGAGCAGACCCGCGAGCCACTGAGTCTCTCCCGCCCGCCG 393
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 145 TyrSer-----HisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 394 TTCACACACAGCCTCCCAACAAAGACCCGCAATGACATCATCTGTGTGAAGATGCA 453
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerSerHisCysProSer 180
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 454 TCGCCAGTCTCCATCACTGCGGTGTGCGACCCCTCACCTCTCTCTCACGCTGTGTCACT 513
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 181 AlaGlyThrLysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPhe 200
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 514 GCTGGCACCAGTCCCTCATTTCCGCTGGGCGAGCAGCTCCAGCCCGCCAGTTACGCGCTG 573
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 574 CCTCACACCTTGCCATCGCCCAACATCACCATCATTCAGCACCAGAGTGTGAGAAGCC 633
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 221 TyrProArgGlnIleAspThrMetPheCysAlaGly---AspLysAlaGlyArgAsp 239
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 634 TACCCCGGCACATCACAGACACCATGTGTGTGCCAGCGTGTGCCAGGAGGGGCGAAGGAC 693
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 240 SerCysGlnGlyAspSerGlyProValValCysAsnGlySerLeuGlnGlyLeuVal 259
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 694 TCTTCCAGGAGTACTCGGGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTATC 753
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 260 SerTrpGlyAspTyrcysAlaArgProAsnArgProGlyValTyrcysLeuAsn 279
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 754 TCCTGGGCGCAGGATCGTGTGCGATCACCGGAAAGCGTGTGTGTGTGTGTGTGTGTGTGC 813
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 280 LysPheThrLysTrpIleGlnGlnThrIleGlnAlaAsn 292
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 814 AATATGTGGACTGGATCCAGGAGACGATGAAGAACAAT 852
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

RESULT 11
US-08-790-137-2
; Sequence 2, Application US/08790137
; Patent No. 5840871
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL PROSTATE-ASSOCIATED
; TITLE OF INVENTION: KALLIKREIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/790.137
FILING DATE: Filed Herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0195 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 833 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-790-137-2

Alignment Scores:
Pred. No.: 1,03e-59 Length: 833
Score: 650.50 Matches: 127
Percent Similarity: 58.62% Conservative: 43
Best Local Similarity: 43.79% Mismatches: 77
Query Match: 40.45% Indels: 43
DB: 2 Gaps: 7

US-09-936-271B-14 (1-293) x US-08-790-137-2 (1-833)

Qy 8 TrpMetTrpValLeuCysAlaLeuIleThrAlaLeuLeuGlyValThrGluHisVal 27
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 30 TGGTTCCTCTGTTCTGTGC-----CTCGCCCTGTCCCTGGGG----- 65
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 28 LeuAlaAsnAsnAspValSerCysAspHisProSerAsnThrValProSerGlySerAsn 47
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 65 ----- 65

Qy 48 GlnAspLeuGlyAlaGlyAlaGlyGluAspAlaArgSerAspSerSerArgIle 67
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 66 -----GGGACTGTGTGTGCGCCCGCGATTGAG-----TCCCGGATT 101
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 68 IleAsnGlySerAspCysAspMetHisThrGlnProTrpGlnAlaLeuLeuArg 87
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 102 GTGGGAGGCTGGGAGTGTGAGCAGCATCCCGAGCCCTGGCAGCGGCT---CTGTACCAG 158
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 88 ProAsnGlnLeuTyrcysGlyAlaValLeuValHisProGlnTrpLeuLeuThrAlaAla 107
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 159 AAGACGCGGCTACTCTGTGGGGCGACGNTCATNGCCCGCAGATGTTCTCTGACAGAGCC 218
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 108 HisCysArgLysLysValPheArgValArgLeuGlyHisTrpSerLeuSerProValTyr 127
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 219 CACTGCTTNAAGCCCGCTCATATGATTCACCTGGGGCAGCACACCTCCAGAG---GAG 275
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 128 GluSerGlyGlnGlnMetPheGlnGlyValLysSerIleProHisProGlyTyrSer--- 146
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 276 GAGGCGTGTGAGCAGACCCCGCAGCAGCAGTCTCTCCCGCCCGCGGCTTCAACAC 335
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 147 -----HisProGlyHisSerAsnAspLeuMetLeuLysLeuAsnArgIle 163
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 336 AGCCTCCCAACAAAGACCCAGCAATGACATCATCTGTGTGTGAAGATGGCATCGCGATC 395
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 164 ArgProThrLysAspValArgProIleAsnValSerSerHisCysProSerAlaGlyThr 183
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 396 TCCATCACCTGGGCTGTGCGACCCCTCACCTCTCTCCCTGCTGTGCTGCTGGGACCC 455
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Qy 184 LysCysLeuValSerGlyTrpGlyThrThrLysSerProGlnValHisPheProLysVal 203
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 456 AGCTGCCTCATTTCCGGCTGGGCGAGCAGCAGTCCAGCCCGCCCGCTTACGCTGCCTCACACC 515
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
```



; EARLIER APPLICATION NUMBER: 60/049,375  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,881  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,880  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,896  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,020  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,876  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,895  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,884  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,894  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,900  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,901  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,892  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,915  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,019  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,970  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,972  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,916  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,373  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,875  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/049,374  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,917  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,949  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,974  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,883  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,897  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,898  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,962  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,963  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,877  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/048,878  
; EARLIER FILING DATE: 1997-06-06  
; EARLIER APPLICATION NUMBER: 60/070,923  
; EARLIER FILING DATE: 1997-12-18  
; EARLIER APPLICATION NUMBER: 60/092,921

; EARLIER FILING DATE: 1998-07-15  
; EARLIER APPLICATION NUMBER: 60/094,657  
; EARLIER FILING DATE: 1998-07-30  
; NUMBER OF SEQ ID NOS: 1227  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 189  
; LENGTH: 1292  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-205-258-189  
  
Alignment Scores:  
Pred. No.: 2,01e-59 Length: 1292  
Score: 650.50 Matches: 118  
Percent Similarity: 68.24% Conservative: 41  
Best Local Similarity: 50.64% Mismatches: 67  
Query Match: 40.45% Indels: 7  
DB: 4 Gaps: 4

US-09-936-271B-14 (1-293) x US-09-205-258-189 (1-1292)

Qy 65 SerArgIleIleAsnGlySerAspCysAspMethHisThrGlnProTrpGlnAlaLeu 84  
Db 264 ACCAGGATCATCAAGGGGTTCCGAGTGCAGCCTCCTCCAGCCCTGGCAGGAGCC--- 320  
Qy 85 LeuLeuArgProAsnGlnLeuTyrCysGlyAlaValLeuValHisProGlnTrpLeu 104  
Db 321 CTGTTCCGAGGAGACGGCGGCTACTCTGTGGGGGAGCGCTCATCGCCCGCCAGATGGCTCCTG 380  
Qy 105 ThrAlaAlaHisCysArgLysLysValPheArgValArgLeuGlyHisTyrSerLeuSer 124  
Db 381 ACAGCAGCCCACTGCTCAAGCCCGCTACATAGTTTACCTTGGGGGAGCAGCACACCTCCAG 440  
Qy 125 ProValTyrGluSerGlyGlnGlnMetPheGlnGlyValLysSerIleProHisProGly 144  
Db 441 AAG---GAGGAGGGCTGTGAGCAGACCCGAGCAGCCTGAGTCTTCCCGCCACCCCGGC 497  
Qy 145 TyrSer-----HisProGlyHisSerAsnAspLeuMetLeuIleLysLeuAsn 160  
Db 498 TTCACCAACAGCCTCCCAACAAAGACCCGCAATGACATCATCTGTTGGTGAAGATGCCA 557  
Qy 161 ArgArgIleArgProThrLysAspValArgProIleAsnValSerHisCysProSer 180  
Db 558 TCGCCAGTCTCCCATCACTGGGTGGTGGCGACCCCTCCTCCTCCTCCTGTTGTCCT 617  
Qy 181 AlaGlyThrLysCysLeuValSerGlyTyrPglyThrThrLysSerProGlnValHisPhe 200  
Db 618 CTGGCACCAGCTGTCATTTTCGGGTGGGGGAGCAGCTCCAGCCCGCCAGTTACGCGCTG 677  
Qy 201 ProLysValLeuGlnCysLeuAsnIleSerValLeuSerGlnLysArgCysGluAspAla 220  
Db 678 CCTCACACCTTGGCATCGCCCAACATCACCATCATTTGAGCAGCAGAGAGTGTGAGAACGCC 737  
Qy 221 TyrProArgGlnIleAspThrMetPheCysAlaGly---AspLysAlaGlyArgAsp 239  
Db 738 TACCCCGGCAACATCACAGACACCATGTTGTGTCGCCAGCGCTGAGGAGGGGGGAGGAGAC 797  
Qy 240 SerCysGlnGlyAspSerGlyProValValCysAsnGlySerLeuGlnGlyLeuVal 259  
Db 798 TCCTGCCAGGTGACTCCGGGGGCCCTCTGTTGTGTATACCAGTCTCTTCAGGAGGATATC 857  
Qy 260 SerTrpGlyAspTyrProCysAlaArgProAsnArgProGlyValTyrThrAsnLeuCys 279  
Db 858 TCCTGGGGCAGGATCGGTGTCGATCACCGGAAAGCCTGGTGTCTACACGAAAGTCTGC 917  
Qy 280 LysPheThrLysTrpIleGlnGlnThrIleGlnAlaAsn 292  
Db 918 AAATATGTGACTGGATCCAGGAGACGATGAAGAACAAT 956

RESULT 14

US-09-025-059-2  
; Sequence 2, Application US/09025059  
; Patent No. 6075136



TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: COLNOT27  
CLONE: 1798496  
SEQUENCE DESCRIPTION: SEQ ID NO: 19 :  
US-09-008-271A-19

Alignment Scores:  
Pred. No.: 1.97e-58 Length: 994  
Score: 639.50 Matches: 126  
Percent Similarity: 57.09% Conservative: 43  
Best Local Similarity: 42.57% Mismatches: 80  
Query Match: 39.77% Indels: 47  
DB: 3 Gaps: 7

US-09-936-271B-14 (1-293) x US-09-008-271A-19 (1-994)

```
Qy 1 MetAlaThrAlaArgPro-----ProTrpMetTrpValLeuCysAlaLeuIle 16
Db 155 ATGGGACGCGCCGACCTCGTGGCGCAAGACGCTGGATGTTCTCG-----199

Qy 17 ThrAlaLeuLeuGlyValThrGluHisValLeuAlaAsnAspValSerCysAsp 36
Db 200 -----CTCTGTGGTGGG-----211

Qy 37 HisProSerAsnThrValProSerGlySerAsnGlnAspLeuGlyAlaGlyAlaGlyGlu 56
Db 212 -----GGAGCTGGCGAGGACAC 229

Qy 57 AspAlaArgSerAspSerSerArgIleAsnGlySerAspCysAspMethHis 76
Db 230 TCCAGGCGACAGAGGAGC-----AAGTGTGGGGGTGATGATGCCAACCCCAT 280

Qy 77 ThrGlnProTrpGlnAlaLeuLeuLeuArgProAsnGlnLeuTyrcysGlyAlaVal 96
Db 281 TCGCAGCCTTGGCAGGCGGCC---TTGTCAGGCGCCAGCACTCTCTGTGGCGGTGC 337

Qy 97 LeuValHisProGlnTrpLeuLeuThrAlaAlaHisCysArgLysLysValPheArgVal 116
Db 338 CTGTGTAGTGCACTGGTCCITTACAGCTGCCACCTGTAAACCCGAAATACACAGTA 397

Qy 117 ArgLeuGlyHisTySerLeuSerProValTyroGluSerGlyGlnGlnMetPheGlnGly 136
Db 398 CGCCTGGGAGACACACAGCCTACAGAAAT---AAAGATGCCCGCAGAGCAAAATACCTGTG 454

Qy 137 ValLysSerIleProHisProGlyTySerHis-----ProGlyHisSerAsnAsp 153
Db 455 GTTCAGTCCATCCACACCCCTGCTACACAGACGCGATGTGGAGGACCAACCATGAT 514

Qy 154 LeuMetLeuIleLysLeuAsnArgArgIleArgProThrLysAspValArgProIleAsn 173
Db 515 CTGATGCTTCTCAACTGGGTGACCGAGCATCCCTGGGGTCCAAAGTGAAGCCCATCAGC 574

Qy 174 ValSerSerHisCysProSerAlaGlyThrLysCysLeuValSerGlyTrpGlyThrThr 193
Db 575 CTGGCAGATCATTCACCCAGCCCTGGCCAGAGTGCACCGCTCTCAGGCTGGGGCAGTGC 634

Qy 194 LysSerProGlnValHisPheProLysValLeuGlnCysLeuAsnIleSerValLeuSer 213
Db 635 ACCAGTCCCGAGAGAATTTCTTGACACTCTCAACTGTGCGAAGTAAATACTTTCCC 694

Qy 214 GlnLysArgCysGluAspAlaTyProArgGlnIleAspThrMetPheCysAlaGly 233
Db 695 CAGAAGAAGTGTGAGGATGCTTACCGGGCGAGATCACAGATGGCTGTGTGCAGGC 754

Qy 234 AspLysAlaGlyArgAspSerCysGlnGlyAspSerGlyGlyProValValCysAsnGly 253
Db 755 AGCAGCAAAAGGGGTGACACCTGCGAGCGGATTTCTGAGGCGCCCTGGTGTGTGATGGT 814

Qy 254 SerLeuGlnGlyLeuValSerTrpGlyAspTyProCysAlaArgProAsnArgProGly 273
Db 815 GCATCTCCAGGGCATACATCCTGGGGCTCAGACCCCTGTGGGGGTGCGACAAACCTGGC 874
```

Qy 274 valtyrThrAsnLeuCysLysPheThrLysTrpIleGlnGluThrIle 289  
Db 875 GTCTATACCAACATCTGCCGCTACCTGGGACTGGATCAAGAAGATCATA 922

Search completed: August 5, 2003, 14:00:12  
Job time : 84 secs